DOCUMENT RESUME

ED 279 609 SP 027 968

| | |
|---|---|
| TITLE | Time for Results: The Governors' 1991 Report on Education. Supporting Works. Task Force on College Quality. |
| INSTITUTION | National Governors' Association, Washington, D.C. |
| PUB DATE | Aug 86 |
| NOTE | 99p.; For summary report, see SP 027 962; for individual task force reports, see SP 027 963-969. |
| AVAILABLE FROM | National Governors' Association Publications Office, 444 N. Capitol St., Washington, DC 20001-1572 ($6.00). |
| PUB TYPE | Reports - Descriptive (141) |
| EDRS PRICE | MF01 Plus Postage. PC Not Available from EDRS. |
| DESCRIPTORS | Accreditation (Institutions); *Change Strategies; *Colleges; College Students; *Educational Assessment; *Educational Change; Higher Education; Outcomes of Education; *State Action |
| IDENTIFIERS | *Time for Results (Governors 1991 Report on Educ) |

ABSTRACT

The Task Force on College Quality was charged with gathering research, reviewing literature and drawing on testimony on assessment of college effectiveness and state policies on improving higher education. Three papers are presented: (1) "College Quality: Measuring What Students Learn in Higher Education" (Jean G. McDonald); (2) "The State Role in Assessing College Outcomes: Policy Outcomes and Probable Impacts" (Peter T. Ewell); and (3) "The Costs of Assessment" (Peter T. Ewell and Dennis Jones). This report includes lists of the charges to the task force, hearings and presenters, highlights of state policies, and additional resources. (JD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

TASK FORCE ON COLLEGE QUALITY

*Supporting Works*

FD279609

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

J. Wills

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

027 968

# Time For Results

## THE GOVERNORS' 1991 REPORT ON EDUCATION

TIME FOR RESULTS: The Governors' 1991 Report on Education

Supporting Works

Task Force on College Quality

CHAIRMAN

John Ashcroft
Governor of Missouri

TASK FORCE MEMBERS

Bob Graham
Governor of Florida
Vice Chairman

George Deukmejian
Governor of California

John Carlin
Governor of Kansas

Harry Hughes
Governor of Maryland

Michael S. Dukakis
Governor of Massachusetts

George Nigh
Governor of Oklahoma

Arch A. Moore Jr.
Governor of West Virginia

August 1986.

Published by the National Governors' Association Center for Policy Research and Analysis, 444 North Capitol St., Washington, D.C. 20001-1572.

Papers in this report were prepared under a subcontract from the National Governors' Association (NGA) Center for Policy Research and Analysis with funding from the Lyndhurst Foundation and the Carnegie Corporation of New York.

Contractors undertaking such projects are encouraged to express their professional judgments. The interpretation and viewpoints stated in this document do not necessarily represent the official position or policy of the NGA or the Lyndhurst Foundation or the Carnegie Corporation of New York and no official endorsement should be inferred. Governmental agencies, groups, and organizations are permitted to duplicate this report in whole or in part, provided that proper credit is cited within the reproduction.

Printed in the United States of America.

TABLE OF CONTENTS

Acknowledgements i

PART ONE: The Task Force

1. Charge to the Task Force 3

2. Task Force Activities 5

PART TWO: Papers

1. College Quality: Measuring What Students Learn in Higher Education
Jean G. McDonald 9

2. The State Role in Assessing College Outcomes: Policy Outcomes and Probable Impacts
Peter T. Ewell 43

3. The Costs of Assessment
Peter T. Ewell and Dennis Jones 75

PART THREE: Additional Resources

1. Highlights of State Policies 105

2. Additional Sources 109

ACKNOWLEDGEMENTS

Many individuals and organizations have devoted enormous time and energy to the year-long effort that created "Time for Results: The Governors' 1991 Report on Education" and the "Supporting Works." The Governors who served on the task forces appointed key staff members who worked not only on their individual task force report, but often assisted other task forces as well. For the Task Force on College Quality these staff members include: Tom Duncan, chairman, Missouri; Debra Gallay, Florida; Carolyn Smith and Scott Ingham, Oklahoma; Sheila Tolliver, Maryland; Gerard Indelicato, Massachusetts; William L. Cunningham, California; Terry Johnsen and Bob Wootton, Kansas; and Catherine Bullard and David Powers, West Virginia. Jennie Carter and Keel Hunt of Gov. Lamar Alexander's office helped coordinate the overall education initiative.

The National Governors' Association Center for Policy Research and Analysis coordinated the year-long project. Professional staff and consultants include Joe Nathan, Joan L. Wills, Evelyn Ganzglass, Carol Eliason, Rae Young Bond, Linda Myers Dove, Jean McDonald, and Peter T. Ewell. Editorial support was provided by Mark Miller, Karen Eldridge, Diane Greene, and Alan Janesch. Expert secretarial assistance was provided by Beverly Bowers, Linda Burnette, Luisa Farrell, Francine Matthews, and Deborah Woods.

The Task Force on College Quality of the National Governors' Association is grateful for the willing cooperation of the postsecondary community throughout task force activities. Organizations representing higher education provided needed assistance as did representatives of colleges and universities, members of state higher education governing boards, and testing services officials. The Office of Educational Research and Improvement (OERI) of the U.S. Department of Education extended assistance whenever called upon. Dr. Charles Haughey, Jim Bencivenga, and Dr. Cliff Adelman, all of OERI, cooperated with the task force in this endeavor.

Special thanks is owed to the following organizations and individuals who helped provide financial support for the task force activities: General Electric Company, Ameribanc (St. Joseph, Mo.), Gerald R. Sprong, West Plains Bank, Norman J. Stupp Foundation, Brown Group, Inc., Union Pacific System, Maritz, Inc., McDonnell Douglas, Pet, Incorporated, Citicorp Person-to-Person, Southwestern Bell Telephone, Stinson, Mag, Fizzell, Emerson Electric, Sverdrup Corporation, Yellow Freight System Foundation, Monsanto Fund, and St. Joe Minerals Corporation.

## Charge to the Task Force on College Quality

The Task Force on College Quality was charged with gathering research, reviewing the literature, and drawing in testimony to address the following questions:

A. What state policies and practices are currently in use to measure the performance of management, faculty, and students in colleges and universities, and how effective have these been in improving quality?

B. Do current measures of student performance accurately reflect community and student expectations?

C. What policies and practices that will measure quality are being developed by the regional accrediting organizations?

D. Are communities and students satisfied with the accuracy and usefulness of the consumer information provided to them regarding course offerings, quality of instruction, and the success of those efforts?

In addition, the task force was charged with answering these policy-related questions:

A. If states choose to develop statewide measures to determine quality outcomes for students, faculty and institutions, what should they be?

B. Should such measures vary by type of institution in accordance with its mission? If so, how can weights be ascribed within the standards set by a state to accommodate these variances?

C. What, if any, interstate measures or indicators are desirable?

D. Should states continue to delegate the same powers as in the past to the regional accreditation organizations? Should there be more state oversight of these organizations?

## TASK FORCE ACTIVITIES

Three hearings were held by the Task Force on College Quality. The first was held in St. Louis, Missouri, December 17, 1985, and witnesses from colleges, universities, and organizations concerned with higher education described models of assessing student learning in college.

St. Louis, Missouri, December 17, 1986

Shaila Aery
Commission
Missouri Coordinating Board for Higher Education

Wayne Brown
Dean, School of Business
Belmont College, Nashville, Tenn.

John Losak
Dean of Institutional Research
Miamia-Dade Community College, Miami, Fla.

J. Terrence Kelly
Vice President, North Campus
Miamia-Dade Community College, Miami, Fla.

Peter Magrath
President, University of Missouri, Columbia
Columbia, Mo.

Charles McClain
President, Northeast Missouri State University
Kirksville, Mo.

Joel Read
President, Alverno College
Milwaukee, Wis.

Frank Newman
President, Education Commission of the States
Denver, Colo.

The second hearing was held in Washington, D.C. February 22, 1986. Witnesses discussed the value and problems of different forms of assessment, methods of assessing student and institutional growth, and quality of today's college students and college education.

Washington, D.C., February 22, 1986

Bruce M. Carnes
Deputy Undersecretary of Education
U.S. Department of Education
Washington., D.C.

Ambassador Stephen Low
Director, Foreign Service Institute
U.S. Department of State
Washington, D.C.

David R. Estes
College Relations Manager
McDonnell Douglas Corporation

Colonel Howard Prince
Professor, U.S. Military Academy
West Point, N.Y.

George Hanford
President, The College Board
New York City, N.Y.

Hoke Smith
President, Towson State University
Towson, Md.

Kenneth H. Ashworth
Commissioner, Coordinating Board
Texas College and University System
Austin, Texas

Richard M. Millard
President, Council on Postsecondary Accreditation
Washington, D.C.

Charles Peters
Editor-in-Chief, Washington Monthly
Washington, D.C.

The final hearing was held in Kansas City, Missouri, May 7, 1986. Participants discussed quality of test instruments available, the role of faculty in developing and using assessment, and the myths of assessment.

Kansas City, Missouri, May 7, 1986

Roxanne Bradshaw
Secretary-Treasurer, National Education Association
Washington, D.C.

Oluf Davidsen
President, The American College Testing Program

Peter T. Ewell
Senior Consultant
National Center for Higher Education Management Systems
Boulder, Colo.

Fred B. Goodwin
Dean, College of Humanities
Southeast Missouri State University
Cape Giradeau, Mo.

Robert Nielsen
Assistant to the President for Higher Education
American Federation of Teachers
Washington, D.C.

Candy Cartwright Young
Assistant Professor
Northeast Missouri State University
Kirksville, Mo.

# COLLEGE QUALITY: MEASURING WHAT STUDENTS LEARN IN HIGHER EDUCATION

Jean G. McDonald
Consultant
National Governors' Association

The 1986 college diploma might bear the label "Warning: This college degree is worth only a fraction of the college degree of 1960," according to recently released reports on postsecondary education. "Evidence of decline and devaluation is everywhere," charges a report by the Association of American Colleges (Turnbull, p. 23). This report and other major reviews of college quality fault the American system of higher education on numerous points. Most worrisome is the contention that "large numbers of students are not able to pass low-level tests of basic skills and knowledge" but still may graduate from college (Southern Regional Education Board, p. 5). Thus, the value of a college degree to a student, future employers, taxpayers, parents, and society is questioned.

These accounts reinforce and supplement assessments of declining quality in elementary and secondary schools, creating severe doubt about the entire educational system from kindergarten through higher education. On the positive side, the reports have encouraged an enormous wave of critical evaluation, discussion, and reform efforts aimed at improving the nation's educational product. Public involvement in this process is substantial, guiding state and local officials and educators in the search for answers to improve American education. One result is greater awareness of the continuity and interrelatedness of the system. Low quality at any stage in the educational process will debilitate education at all levels. For example, inadequate preparation in elementary school affects performance in secondary schools; deficient training in secondary schools results in extensive remediation in college; weak college teacher preparation programs impair instruction at the elementary and secondary levels.

As the prime source of funds for public higher education, the states have a major stake in the quality of postsecondary institutions. Traditionally, college quality has been measured by the accreditation process. In addition, quality has been judged by measurements such as faculty/student ratio, faculty research publications, number of library books, research equipment, caliber of student applicants, ability to place graduates in jobs and graduate schools, and alumni success. Seldom has student learning been considered as a factor in college quality. No one has known how much students learn except through surrogate measures such as grade point average, acceptance to medical school, and failure or passage of licensing exams. In fact, Jonathan Warren has argued, we do not know if today's education is inferior to what it has been in the past because we lack hard data about what students have learned in the past. "Judgments about the quality of American education are based on little more than folklore, nostalgic reminiscences and personal preferences" (Warren, 1981, p. 5).

---

This paper reviews relevant literature and research on the assessment of college students, distills the testimony presented at the three hearings on college quality and assessment, and presents models of state activities and programs in college assessment. The author owes a special debt of gratitude to Dr. Cliff Adelman, Office of Educational Research Improvement, U.S. Department of Education, Dr. Ted Marchese, Vice President, American Association of Higher Education, and the many officials in state higher education governance offices who described programs in their state.

Past failure to measure what students have learned in college will not be accepted as an excuse for present inaction by the many champions of reform. The reform reports view assessment as an integral part of efforts to improve college quality: "Faculty and academic deans should design and implement a systematic program to assess the knowledge, capacities, and skills developed in students by academic and co-curricular programs," recommends the NIE report, Involvement in Learning. Governors and legislators who fund and oversee public systems of higher education are asking for better measures of student performance. Employers, disappointed in graduates' communication skills question the meaning of a college degree (Low, Estes). Parents, slapped with higher bills for college, want to know what their children are getting out of their expensive education. Taxpayers, concerned with the best possible use of tax dollars, want public education, at all levels, to be held accountable. As Missouri Gov. John Ashcroft, task force chairman, stated at the second hearing, "while the faculty and publicly funded institutions need to be free to teach freely, the public has the right to expect that substantive learning does take place; public officials have the responsibility to ensure that learning indeed occurs."

Lower standards, watered-down curricula, "fad courses," and declining performance on standardized examinations, all of which accelerated in the late sixties and seventies, contribute impetus to the college quality issue, in general, and to the assessment issue, in particular. For example, in 1983 only 50 percent of all postsecondary institutions required foreign language studies for the bachelor's degree, down from 90 percent in 1966 (Carnes). An American Council of Education survey in 1984-85 discovered 28 percent of American colleges offering baccalaureate degrees do not require American literature and American history (Carnes). Average scores on fifteen of twenty-three standardized examinations used in selecting students for graduate school admission have declined since the mid-1960s (Adelman 1984). The Graduate Record Exam, both general and specific subject matter, the Law School Admissions Test, and the Graduate Management Admissions Test registered sizable declines in average student performance in the 1964-1970 period, leveled off in 1970-1976, and dipped slightly from 1976-1982.

A closer examination of the scores reveals the declines are unrelated to changes in the contemporary college student population on factors such as age, race, gender, and citizenship. Rather, the most apparent explanation for declining test scores is the greater number of students choosing professional and occupational college majors (engineering majors are an exception). Fewer of today's students select majors in traditional arts and sciences fields, but these students score higher on the standardized exams. "Students with undergraduate majors in science, mathematics, and engineering perform better than all others on these examinations" (Adelman 1984, p.23).

Postsecondary institutions have begun to respond to the outcries for reform, improved standards, and assessment of student learning. An American Council of Education (ACE) 1984 survey revealed four of ten postsecondary institutions had established writing requirements, two-thirds required placement tests, and 62 percent offered courses aimed at retaining students. ACE's 1986 survey results reveal that 86 percent of American colleges are making curricular revisions or have already done so. Impressive support for

assessment was evidenced by the most recent survey: 91 percent of the respondents favor assessment programs linked to instructional improvement (El-Khawas 1986). Schools such as Northeast Missouri State University, State University of New York at Plattsburgh, Kean College (New Jersey), Alverno College, University of Arizona, and City University of New York are all examples of colleges and universities that have adopted programs assessing student outcomes.

## CHANGES IN POSTSECONDARY EDUCATION

The wave of interest in upgrading postsecondary education and assessing student learning arrives at a time colleges and universities are facing challenges from an increasingly diverse and occupationally oriented student body. The traditional eighteen- to twenty-four-year-old, full-time, white, middle-class student is increasingly less visible on college campuses as demographic changes, already experienced at the elementary and secondary level, work through the pipeline. The number of eighteen-year-olds is projected to decline through 1995 when a new "baby boomlet" will reach college age. For example, in 1985 eighteen-year-olds numbered 3.7 million but by 1995, the eighteen-year-old population is expected to drop to 3.3 million. And by 2000, this age group is projected to reach 3.8 million (U.S. Bureau of the Census). The percentage of college students twenty-five years old or over has been on the rise since 1974: 36 percent of college and university students in 1984 were twenty-five or older (U.S. Bureau of the Census) and of these, 56 percent were female (U.S. Bureau of the Census). Arts and sciences degrees have dropped from 49 percent in 1971 to 36 percent in 1982, while occupational and professional degrees have risen proportionately (National Institute of Education).

Nontraditional students are more likely to enroll part time and take longer for completion. Part-time students represented 42 percent of total college enrollment in 1982, up from 32 percent in 1970 (National Center for Education Statistics, 1984, p. 70). A recent nationwide survey of 5,000 undergraduates by the Carnegie Foundation for the Advancement of Teaching confirmed the description of part-time students as nontraditional. Three-fourths were employed (one-half on a full-time basis) and "they were generally older, more than half of them were married, and about one-third had children" (Jacobson, p. 30). However, even traditional students may take longer than the usual four years to complete a degree as they take time out for financial and personal reasons. Half of today's students complete a college education in four years; 70 percent complete degrees within seven years of beginning their program (Hodgkinson).

Minority student enrollment patterns are related to assessment and reform issues; nonwhite students may be hard hit by increasing standards and requirements. It is well documented that many minority and poor students perform less well on standardized achievement tests and are less likely to be enrolled in college preparatory high school programs than middle-class white students (Navarro). In 1984, minority students comprised approximately 20 percent of the 1984 college student population (Office of Minority Concerns). The Condition of Education reports minority enrollment in 1982 at 14 percent in four-year institutions and 20 percent in two-year postsecondary

institutions (National Center for Education Statistics, 1984, p. 63.) Given projected population figures, the proportion of minority students will expand at the K-12 level in the future, but that *growth may or may not result in* higher minority college enrollments. Historically, smaller proportions of minorities than whites attend college. In 1984, 34 percent of white high school graduates age eighteen to twenty-four were enrolled in college, compared with 27 percent of black peers and 30 percent of the Hispanic cohort (U.S. Bureau of the Census).

Higher percentages of minority high school graduates do not necessarily translate into larger numbers of minority college students. While Hispanic, native American, and Asian enrollments have been on the rise since 1978, total black college enrollment has remained stable (Office of Minority Concerns). Yet the pool of eligible black students is increasing: the percent of eighteen- to twenty-four-year-old black high school graduates rose from 65 percent in 1979 to 75 percent in 1984 (U.S. Bureau of the Census). Obviously, access to higher education is not yet a reality for many qualified minority students.

Remedial needs are great at today's colleges and universities as more entering students, regardless of age or race, require assistance in basic skills. According to Hodgkinson, 80 percent of postsecondary institutions are offering remedial course work. Eighty-eight percent of colleges surveyed in 1986 require placement tests in either math or English, the results of which may be used to place students in remedial classes (El-Khawas 1985). Such courses may or may not be offered for credit, depending on institutional or state policy.

Postsecondary institutions also are affected by changes in faculty. Faculty are leaving the teaching profession for better paying jobs outside academe. In combination with the smaller proportion of Ph.D.s being attracted to college teaching, institutions of higher education may face a shortage of faculty in the near future (Bowen and Schuster). A 1984 survey of 5,000 faculty by the Carnegie Foundation of the Advancement of Teaching found an alarming degree of faculty dissatisfaction with higher education. "In all ranks, the prevailing faculty view is that academic standards are too low, students are too weak, the campus intellectual environment is only fair at best, and faculty compensation and working considerations are inadequate" (Jacobson, 1985, p. 1). A rise in part-time faculty (from 28 percent of total senior instructional staff in 1970 to 53 percent in 1982: National Center for Education Statistics, 1984, p. 78) has occurred to offset faculty shortages and funding problems. This is not viewed as a positive development. David Estes of McDonnell Douglas Corporation testified at the second College Quality hearing that "the biggest concern that we see facing our company and other high tech companies is the retention of faculty." Enrollments in essential courses are limited because faculty are lured to more lucrative jobs outside the universities. Faculty salaries have not kept up with salaries in other sectors; in 1985-86, average faculty salaries ranged from $26,240 for assistant professors to $42,500 for full professors, across all institutions, according to the American Association of University Professors.

All of these trends -- an older, nontraditional, and increasingly part-time student body, increased minority participation (with the exception of black students), greater need for remedial course work for entering students, and faculty disenchantment with higher education -- have crucial implications for postsecondary institutions. Colleges and universities must be flexible in meeting the challenges posed by these changes at the same time society expects them to enact higher standards. Assessment practices, in combination with remediation and effective feedback to students, may result in greater public confidence and, ultimately, may strengthen the position of the higher education community.

## HISTORY OF ASSESSMENT

Assessment refers to a comprehensive process of testing and evaluation for individual progress in learning and ability to perform. When the term was first used by psychologists in the 1930s and 1940s it was based on the assumption "that an individual's performance and potential was too broad to address with single measures" (Marchese). Performance, testing, and feedback were basic principles tested at the Harvard Psychological Clinic under Henry A. Murray in the late 1930s. The theory was put into practice by the Office of Strategic Services during World War II. In the 1950s AT&T effectively adopted the model to identify managerial strengths, and by the 1960s and 1970s assessment programs were common in the corporate sector (Hartle).

Testing for performance and overall learning is not new to the college campus. Comprehensive assessment was tried early in the Hutchens College at the University of Chicago and the General College at the University of Minnesota (Ewell). Pennsylvania had a large-scale testing effort from 1928 to 1932 through the General College Test Program. Liberal arts schools have used comprehensive senior exams to determine what students learn in college (Ewell). The Swarthmore comprehensive exam, developed in 1922 as part of the honors program, included four exams in the student's major and two in the minor. Faculty from other institutions, within the student's discipline, served as external evaluators in the grading process (Harris).

In the late 1960s Alverno College, a small private liberal arts college in Milwaukee, undertook a curriculum revision with an eye toward the end product of college, e.g., the desired qualities, abilities, and skills of Alverno graduates. This task required consideration of how to progress from point A, the incoming student, to point B, the graduating senior. In searching for model programs to emulate, the Alverno faculty turned to the business community, lacking higher education models. AT&T's assessment program, using multiple measures and continual feedback, was modified to college program needs. Today, the Alverno assessment process is itself considered a model program and will be described in more detail below.

When Northeast Missouri State University (NMSU) underwent the transition from a teachers college to a "small multipurpose university" in the early 1970s, a decision was made to adopt a program of assessment of student learning, known as value-added assessment. Scores from tests administered to entering college freshmen were compared with identical tests given to second semester sophomores, thereby demonstrating gains from the college experience.

Exit testing, demographic data, and student and alumni attitudinal surveys round out information gathered to show the value of an NMSU education. As with Alverno, the NMSU program has undergone numerous revisions through the years and has become a model of a value-added assessment program (McClain; McClain and Krueger).

Other relatively early attempts to assess what students learn in college occurred in a variety of higher education settings. Our Lady of the Lake University in Texas, which serves a nontraditional student body (53 percent Hispanic, 69 percent female, 65 percent over 21, and 48 percent part time) has utilized an assessment program to determine the outcomes of the General Education program since the early 1970s. Lander College (South Carolina) became an public open admissions college in 1973 after a history as a private school. To adjust to its changing mission, placement testing for all incoming freshman was instituted in conjunction with remedial course work. The Regents External Degree Program (New York) was initiated in 1970 in an effort to provide an "examination university" for nontraditional students. While no classes are offered by the institution, adults who demonstrate proficiency of subject matter through written and performance examinations may obtain degrees in eight fields (Adelman 1984). In 1973, Georgia began its "Regent Exam," requiring college students to demonstrate English proficiency before graduating. These initiatives, and many other recently adopted assessment programs described more fully below, have provided the momentum to place the issue of assessing what students learn in college squarely on the Governors' agenda in 1986.

## COMPONENTS OF ASSESSMENT

While teacher-developed tests and grades have been the traditional method of assessment in colleges and universities, the assessment movements calls for postsecondary institutions to move beyond grades to a more systematic evaluation of the overall college experience. No one is advocating the elimination of grading systems or faculty autonomy in the classroom, but policymakers and others are calling for less subjective tools of measurement and for measures that can compare students within a university as well as with students at other universities. Grades are a limited mechanism to assess the total student performance or institutional performance in delivering educational services. In addition, grade inflation has been a recent trend -- it appears students are receiving higher grades for lower performances. Grade inflation is possible when grades are curved to the individual classroom, standardized measures of class performance are lacking, and professors lack training in evaluation methods.

Terry Hartle, writing for the 1985 conference on assessment, uses The Encyclopedia of Education Evaluation's definition of assessment as a "multi-trait, multi-method approach" (Hartle, p. 4). Hartle classifies six types of assessment: 1) multiple measures and observations used to track student growth over time; 2) state mandated requirements established to evaluate students, academic programs, and universities; 3) value-added assessment; 4) general standardized testing used for assessment purposes; 5) assessment practices used as part of the formula for institutional funding; and 6) surveys to measure the influence of college years on students' attitudes and values.

Theodore Marchese, vice president of the American Association of Higher Education, states that assessment is an attempt "to sort, screen, qualify, or certify students; or to measure the performance of an institution or programs within it" (Marchese, p. 10). Concerned that assessment not rely on a single score, Marchese suggests a multitude of achievement and performance measures available to measure achievement. Direct observation of performance, expert opinion, standardized achievement tests, peer judgments, diagnostic batteries, structure interviews, self-assessment, and review of performance on written exams from classes are suggested tools of evaluation.

Marchese is not alone in arguing on behalf of multiple indicators to assess both student and institutional performance (Ewell, Hanford, Hoke Smith, Warren). The diversity of the American postsecondary institutions is one reason multiple measures should be used, they assert. Measurement must be sensitive to an institution's mission, e.g., an open admission two-year college emphasizing occupation and vocational training versus a four-year baccalaureate program, versus the liberal arts institution versus the large research university, and so on. Laurence Marcus, Anita Leone, and Edward Goldberg refer to a New York State Department of Education report (1981) that recommends that assessment be multidimensional and account for the institution's mission. Warren discusses a number of measures of institutional quality, including the aggregation of course information from transcripts, catalogs, syllabi, and in-class exams. The Maryland State Board for Higher Education report "Alternatives for Testing Maryland College Students" (1984) suggests "It is important that testing be viewed as only one of many possible means by which to evaluate campuses." Employer surveys, alumni studies, and outside expert evaluation are also recommended in the report.

Writing in 1969, Alexander Astin and Robert Paros developed a system's model of assessment emphasizing the input, environmental, and output dimensions of assessment. Input refers to the academic skills, self-discipline, attitudes, and characteristics that entering freshmen bring to college. The environmental variable is the college experience, both academic and nonacademic. Outputs have three dimensions, according to Astin and Paros: cognitive (knowledge) versus affective (attitudinal) factors; the type of data used in assessment (psychological versus behavioral); and long-term versus short-term effects. Ewell furthers the understanding of assessment by his typology of four dimensions: cognitive development, skills development, attitudinal development, and behavior after college.

Two-year colleges have a stake in documenting program quality to assure that degrees mean what they purport to mean (ERIC Digest). Four methods have been used by community colleges to demonstrate the commitment to quality assurance. Mastery learning programs have been adopted, requiring documented subject matter or skill proficiency. Competency-based education programs have been instituted, requiring proof of student ability to perform specific skills and behaviors. Curriculum tracking or successively ordering classes assists students in completing their degree program. Finally, minimum competency tests at entrance and exit demonstrate students educational progress.

No discussion of college and university assessment can be complete without reference to accreditation, the traditional method of tracking college quality in America. However, even accreditation programs are beginning to note the need for substantiation of student learning progress. The Commission on Colleges of the Southern Associations of Colleges and Schools has adopted standards for accreditation that include an assessment component. Member institutions are expected to establish educational goals, demonstrate how the goals are being attained, and document that the goals are being reached. The commission has been careful to indicate the importance of multiple measures of assessment, beyond quantification of test scores.

## Advantages of Assessment

At the second College Quality Task Force hearing, Gov. Ashcroft listed several positive benefits of assessment. "Assessment of student learning, when done properly, with appropriate instruments and with trained assessors, is a useful approach for the demonstration of instructional effectiveness, for the assessment of quality in program and curriculum, for demonstrating accountability to students, to parents, to taxpayers and, of course, to sponsoring institutions like state government." Other advantages are evident when operating assessment programs are examined. Increases in student learning have been documented at Northeast Missouri State University. Curricular changes often accompany assessment programs as faculty and staff utilize test and survey information to pare "fad" classes and strengthen the curriculum. At the University of Tennessee, Knoxville, departments responded to test and student data by changing introductory courses and curricular structures (Ewell). Goal development, an essential first step in assessment, helps faculty clarify expectations of student learning (Harris).

Assessment places renewed emphasis on undergraduate education and the teaching and learning efforts that go into that process. Public satisfaction appears high with evidence of accountability, and policymakers can point to concrete proof of student learning. Assessment practices accompanied by funding incentives, such as the Tennessee system and rewards to NMSU, enhance the acceptance of a campus assessment program. Also, assessment may encourage greater interaction and cooperation between secondary schools and institutions of higher education. Southeast Missouri State University developed an English proficiency exit exam in 1984; faculty members have worked with high school English teacher to show high schools students the expected level of English proficiency in college. Carnes, testifying at the second task force hearing, advocated truth-in-advertising standards for postsecondary institutions; schools would be required to publish information such as student retention and completion rates and placement rates. Such information may be easily supplied by colleges with comprehensive assessment programs.

## Criticisms of Assessment

Despite the many benefits of assessment programs, doubt, cynicism, and fear have been expressed by the academic community. "You cannot subtract what someone knows at the beginning of college from what he knows at the end and think that the difference is what he learned. Education is not simply the piling up of pieces of information." Jonathan Warren's statement is typical

of those who contend that the broad range of learning and higher order skills characteristic of a college education cannot easily be subjected to a single test. And basic skills testing is too narrow a goal for a college education, confusing the mission of college with that of public schools (Turnbull). Marchese argues that assessing student performance is a more difficult but ultimately a far more essential task than assessing student learning. Professions, business, and the public want to know how well the teacher or physician performs, not just what they know.

The notion that assessment practices will lead to "teaching to the test" is a universal criticism. Some evidence for teaching to the test is seen in courses to help students prepare for the national graduate school admissions exams. Miamia-Dade Community College officials admit that the CLAST exam, required of all candidates for degrees, has narrowed the curriculum extraordinarily in order to improve student passage rates. Other concerns about assessment have been expressed. Declines in minority participation and the access issue is of concern to many. This will be discussed in greater detail below. George Hanford of the College Board sees a danger in assessment if policymakers apply the same standards to higher education as have been applied to secondary schools. In addition, "forcing colleges and universities into a more common academic mold (could) have the effect of eroding one of the major strengths of our American higher education system" (Hanford). Dr. Robert Nielsen of the American College Testing program reiterated this concern at the third hearing by posing a question: how to evaluate to a common standard and yet maintain diversity and access, excellence and creativity. Marchese expressed additional concerns regarding assessment: subsuming other objectives for those measured by the exam; rigid screening and placement decisions affecting individual students; the need for examinations used in assessment to be psychometrically correct (Marchese 1986).

The need for faculty involvement at the center of assessment efforts is strongly recommended for successful programs. Interestingly, however, the Northeast Missouri State University program was initially proposed by the administration over some faculty opposition. Through the years the NMSU faculty have been increasingly involved in and pleased with the value-added program, observing the many positive benefits to the institution (McClain, Young).

## Instrument-Related Issues

Warren suggests that "most people see the lack of good evaluation instruments as a major impediment to moving ahead (with assessment programs)." Clearly, appropriate testing -- using tests that accurately evaluate the desired outcomes -- is extremely important in a legitimate assessment effort. Concern regarding the proper test to use is not misplaced as the stakes are too high to treat the question lightly, particularly if students are to be denied promotion or diplomas on the basis of an examination score, as in Florida or Georgia.

The recent Congressional Budget Office report "Trends in Educational Achievement" stresses that tests must be valid and reliable. Examinations are valid if they test the skills for which the test is intended; curricular

validity is "the closeness of the fit between the test and the curriculum to which students are exposed" (Koretz, p. 21). Litigation over public school minimum competency has challenged tests on grounds of a weak match with the curriculum. The reliability of a test refers to the consistency of its results in one administration after another. A test that is unreliable is also invalid.

Another issue in educational testing is whether the test is norm-referenced or criterion-referenced. A norm-referenced test compares the performance of students to other students; only questions that discriminate between students are included in a norm-referenced examination. Questions answered either correctly or incorrectly by 100 percent of the test-takers are eliminated from the test in order to obtain scores that approximate a bell-shaped curve. Harris argues that since norm-referenced tests discriminate between individual students, they emphasize native ability more than knowledge from course work. Such tests are, as a rule, renormed every seven years, which may include revising tests and testing with a fresh sample (Harris).

Criterion-referenced examinations test student ability to reach an absolute standard of performance, as in minimum competency tests. An example of a criterion-referenced test is provided by testing of pilots in a flight simulator. Either the student pilot can reach the standard or he can't. To determine how well student can write, a criterion-referenced test is most appropriate. The Georgia Regents Test for writing proficiency is criterion-referenced, using trained faculty with experience in judging examinations against preset standards. Criterion-referenced tests allow comparisons in knowledge gains for individual students, as well as comparisons of instructional approaches (Harris).

Several other test-related issues have been addressed by both proponents and opponents of assessment practices. The fit between a department's course offerings and a standardized exit exam, such as the GRE, is one issue. For example, the GRE history exam omits questions on Asian and Latin American history, although departments using the exam may want to emphasize those subject areas; student scores could suffer. Further, the university, department, and public or state agency must decide whether to test for academic knowledge or performance or both. Performance requires different assessment measures than knowledge. Colleges and universities must acknowledge that differences in instructional quality may bear some responsibility for differences in student performance. The "stakes" issue is another concern; if students are not taking the exam in order to graduate from college or in order to be admitted to graduate or professional school, they may lack a stake in its outcome. Accordingly, low student motivation may depress performance, yet the "stake" is high for departments and institutions whose effectiveness is judged by the students' scores (Khourry, Taylor, Maryland State Board for Higher Education).

As an alternative to using standardized exams, departmental faculty may decide to develop their test either in general education or in the major field. Although extremely time-consuming, a department may feel it is worth the effort. Trudy Banta at the University of Tennessee, Knoxville, assists

faculty in departments to develop tests assessing major field proficiency if suitable national tests are not available (Harris).

## Access: Will Assessment Hurt Minority Students?

Traditionally, minority students perform less well on standardized exams than do white students. Implementing standardized tests, particularly tests serving as "gatekeepers" whose failure keeps a student from graduating, may serve to keep fewer minority students from higher education opportunities. Yet, there is general agreement that access to higher education must remain open, presenting a major dilemma (Hartle). George Hanford of the College Board, concerned about narrowing minority opportunities for higher education, urged the task force "to avoid excessive externally mandated examinations, and, instead, encourage institutions to embark on assessment themselves. No single instrument can effectively assess academic achievement across all institutions."

Since minority students disproportionately enroll in community colleges, those institutions must be particularly sensitive to access issues. Terry Kelley, vice president, and John Losak, dean of institutional research at Miamia-Dade Community College, testified at the first college quality hearing, that minority students at that institution are being affected by the CLAST exam requirement for graduation. Of 65,000 students enrolled for credit, 48 percent are Hispanic, 16 percent black, 33 percent anglo-white, and 3 percent other. Since 1984, when the exam was first administered, the number of graduates is down by 25 percent. In 1984, 88 percent of anglo students and 62 percent of black students passed. However, as standards become increasingly strict for the exam, peaking in 1989, Kelley and Losak predict that about 25 percent of the student body will pass the exam at present levels of performance; only 14 percent of black students can be expected to pass. The college has revised its curriculum, offering courses focused on skills needed to pass the exam, and great efforts are being made to assist students in all possible ways. Dr. Robert McCabe, president of Miamia-Dade Community College, fears the present testing system will exclude minorities from educational opportunity. McCabe advocates multiple assessment rather than a single score evaluation process.

## Remediation

The need for remedial offerings at contemporary colleges is well documented, and institutions are responding to this need. Remediation is another side of the access issue; however, regardless of race, today's students need remediation to be brought up to standards before and during college (Hartle). Entry-level exams are administered by many institutions and, in some states, such exams are required in order to place students in appropriate remedial classes. Course credit may or may not be given for remedial work. According to the Education Commission of the States, nearly every state has taken action to improve information on college preparation for high school students and strategies to reduce the need for remediation.

Retention of students in college is a related issue. College student dropout rates are high, particularly for minority students. The first

semester of the freshman year, in particular, is a time when many students are discouraged by the college experience and decide not to continue. Increasing retention rates is viewed not only as a way to help students, but also as a mechanism for offsetting population declines in the eighteen- to twenty-four-year-old age group. Additional remediation classes, workshops and courses in study skills and test taking, library proficiency, intensive counseling, and increased financial aid programs are a few of the strategies postsecondary institutions may adopt to lower the dropout rate.

New Jersey has developed a special entry-level test for placement purposes, the New Jersey Basic Skills Placement Test. An evaluative component to this program provides excellent data on the performance of college students at entrance and the impact of remediation. Analysis shows retention rates at both two and four semesters are higher for remediated students than for students who did not need remediation coursework (New Jersey Basic Skills Council 1985).

## STATE POLICY OPTIONS FOR ASSESSING STUDENT PERFORMANCE

A wide range of policy options is available for state policymakers interested in accountability for higher education through assessment. Policies range along a continuum from state mandates requiring all public institutions (and private schools receiving state aid) to institute assessment practices, at one end. Florida represents an example of this type of policymaking, while Michigan represents the decentralized, deregulated policy type. Most states fall somewhere in between, such as Virginia, which leans toward the autonomy model. Virginia state colleges and universities are asked to adopt assessment programs and to report results annually; the type of program is left up to the institution.

Mandates are considered a less satisfactory alternative, according to spokespersons for the academic community and the public sector. The states have a number of means other than mandates to encourage institutional assessment practices. "Funding mechanisms, regulations and reporting requirements, admissions standards, financial aid and other requirements, admissions standards, financial aid and other access mechanisms, and program approval and review policies" can stimulate institutional change (Ewell), while remaining sensitive to the diversity of mission of the public higher education system. John Wallace of the University of Minnesota suggests increasing the use of pilot programs of assessment. Hoke Smith, president of Towson State University, told the task force that the states must ensure that each institution has the staff and computer facilities adequate to the assessment task. States may establish funds to support changes and innovations resulting from the assessment process. Finally, he suggests states should request that accrediting associations review institutional assessment systems as one aspect of the accreditation process.

## INSTRUMENTS OF ASSESSMENT

Existing data, as Peter Ewell and Dennis Jones suggest, may provide essential information about student outcomes and attitudes. Professors' descriptions of student work and comparative analyses of present with former

students, transcript analysis, samples of students' written work, voting in student elections, and other material could supplement a wide array of existing materials which are lying on the shelves unanalyzed. Harris suggests seniors be required to complete a major product, such as a senior thesis, which could be evaluated by outside faculty or experts. Warren suggests mutually developed examination questions that may be used in several sections of the same course or across institutions. William Turnbull emphasizes the importance of collecting additional indicators of institutional functioning such as dropout rates, use of library facilities, and success rates in graduate school entry.

Entry-level tests may be administered to students prior to college entrance either for placement purposes or for college admission purposes. The New Jersey Basic Skills Placement Test, developed in 1977, to test reading, writing, computation, and elementary algebra skills of incoming freshmen is one example of the former. Traditional standardized achievement tests such as the Scholastic Aptitude Test (SAT) and the American College Testing Program (ACT) represent examples of the latter. Another example of a placement test is the Multiple Assessment Programs and Services test used by Florida, Tennessee, and California. Locally developed placement tests are common at postsecondary institutions. The Educational Testing Service (ETS) has developed computerized entry level skills assessment exams that tailor the difficulty of the test to the skills of individual examinees. According to ETS, the computerized exams are more efficient across a broad range of skills than conventional paper and pencil tests (Doyle).

Some tests are classified as "rising junior" exams, a not altogether accurate depiction. In some instances, such as Florida, the exam must be passed prior to entrance to upper division course work or an associates degree. The College Level Academic Skills Test, developed by Florida faculty, examines communication skills in reading, writing, listening, and speaking in addition to mathematics skills. The test is "secured" in Florida and may not be used outside the state (Harris). New Jersey is in the process of developing an exam to be administered at the end of the sophomore year; however, the exam's purpose will be one of information gathering and value-added assessment rather than gatekeeping.

Exams to demonstrate proficiency in writing have been used in American universities for many years. Georgia uses the criterion-referenced Regents Testing Program to ensure that graduates of the university system in Georgia demonstrate English proficiency. The test contains both writing and reading components; the essays are graded by selected faculty (Harris). The California State University System requires evidence of English proficiency prior to graduation; however, individual campuses to develop their own proficiency measures.

Tests to evaluate the outcomes of general education courses have, in some instances, been used to determine the value-added element of a student's first two years of college. The ETS Undergraduate Assessment Program tests in three general areas, humanities, social science and natural sciences. The exam has not been renormed for several years, limiting its usefulness (Marchese). ETS will loan the tests to colleges and universities on a yearly basis but the

instruments must be scored locally. Thus, the test information is useful only to the campus and does not allow comparison with a national sample (Harris). The Academic Competencies and General Education exam was developed by Jonathan Warren for ETS with faculty-generated questions, field-tested on students. Higher order skills, analytic and synthesizing skills, and ability to perceive differentially all are tested by this exam. At the present time, this test is not being utilized. One can expect that in this, and other areas, the major testing areas will develop appropriate tests as assessment practices become more widespread.

The ACT College Outcomes Measures Program (COMP) is an increasingly popular measure of general education. The test has three parts that may be used together or separately -- an objective test, a composite exam, and the COMP Activity Inventory. Both the composite and objective examinations cover three "process" areas including "oral and written communication, problem solving, and values clarification. The three content areas are 'functioning within social institutions,' 'using science and technology,' and 'using the arts'" (Harris, p. 16). Students are asked to respond to audiovisual stimuli, such as a film clip, to analyze a magazine article or advertisement or other similar activities. The entire test battery takes about a day; consequently, it is less common to administer all three test components.

The College Level Examination Program (CLEP) was originally developed for adult students returning to college to demonstrate subject matter proficiency for course credit. Today the exams are frequently used by eighteen-year-olds to test out of basic courses. In addition to a number of subject matter specific exams, CLEP offers six general exams, nationally normed, which are used by some institutions for general education assessment purposes. Other tests, such as PEPS (designed by ACT for the New State Department of Education) and DANTES (designed by the military for evaluation of course mastery) are similar types of examinations.

Senior-level tests may be general or discipline-specific; in the latter case, the exam tests knowledge and competence in the major field. Senior exams, such as written comprehensive examinations, may be used as gatekeepers to graduation. In general, however, senior exams are administered to provide information regarding proficiency in the major field or to be used as an element in graduate school admission. Licensing exams are subject matter-specific, requiring demonstrated proficiency in a professional area, such as architecture. Students must obtain passing scores in order to be admitted into the examined profession.

The Graduate Record Examinations are used to test either general or specific subject matter knowledge. The general exam tests verbal and quantitative skills of senior-level students while the subject matter tests require more specialized knowledge in areas such as biology, chemistry, psychology, political science, and Spanish. These tests are norm-referenced and therefore are poor indicators of program quality (Harris). However, the exams are used at times to compare student scores locally with a national sample. In this indirect manner the test "measures" program quality. In addition, the subject area exams are deemed inappropriate "exit" tests for all seniors in a major because they "measure achievement in those areas that are

considered important for success in graduate school" (Maryland State Board for Higher Education, p. 3). Other graduate school admissions exams, LSAT (law school), MCAT (medical school), and GMAT (management school) are used as indicators of senior-level proficiency.

Performance measures have traditionally been used to assess proficiency in a number of academic disciplines. Architectural education has long relied on juried assessments of student work. Fine arts programs require performance-based assessment. Peer juries, portfolios, faculty evaluation, and outside expert assessment, in addition to self-assessment, may be used to measure proficiency. Sandra Elman and Ernest Lynton advocate performance-based assessment for career-oriented education in a forthcoming chapter in the OERI publication on assessment. Small group interaction, observations of student performance, and expert judgment of student performance all are recommended by the authors. In particular, Elman and Lynton cite the American Management Association's assessment model used to analyze the performance of outstanding managers as a way to assess student competence. Simulations of managerial situations, videotaped exercises, recorded interviews, and a battery of tests produce a student competency profile.

Student and alumni surveys tap attitudes regarding campus quality, faculty, and facilities. Invaluable information regarding the effects of college on student values and "good neighbor" qualities may be elicited by a variety of survey instruments. Attitudinal surveys are locally produced or standardized. Examples of the latter include eleven ACT surveys ranging from Adult Learner Needs to Postsecondary Plans; eight ETS surveys for institutional use, such as Student Reaction to College or Program Self-Assessment Service; Pace's College Student Experiences questionnaire; the National Council on Higher Education Management Systems' (NCHEMS) Student Outcomes Information Service (SOIS); and ACE/UCLA Cooperative Institutional Research Program (CIRP). Lord Fairfax and Piedmont community colleges in Virginia have conducted follow-up studies of the success of former students at George Mason University and the University of Virginia as one measurement of achievement (Report of Study Conducted by the Council of Higher Education for Virginia).

## MODELS OF ASSESSMENT

### Blue Ribbon Study Commissions

Twenty-four states have responded to calls for reform through the establishment of blue ribbon commission on college quality (Edgerton and Marchese). Some of these efforts have concentrated specifically on assessment; others have addressed assessment tangentially as one of many reform issues. In Illinois, the Committee on the Study of Undergraduate Education of the Board of Higher Education is recommending student assessment programs be instituted to provide information on student performance, but will not mandate the details of the assessment package. The Arkansas Board of Higher Education has asked individual institutions to review, modify, and implement campus assessment plans and report to the board on the specific assessment program initiated on campus.

The Louisiana Board of Regents examined the core curricula of state institutions in its general education review. In addition to mandating specific course requirements for general education purposes, the regents ask state public colleges and universities to assess the outcomes of the general education curriculum. However, the decision as to type and form of testing is left to the individual institutions. In June 1985, New Jersey established an advisory commission to design and implement a long-range College Outcomes Evaluation Project. Outcomes were painted with a broad brush to embrace student development, research, scholarship, and contributions to community and society. An instrument to test for these student outcomes is planned for development by 1988. A working group appointed by the Maryland State Board for Higher Education to study assessment options has recommended a general education proficiency test be developed. The exam would be administered to a sample of students in all statewide public postsecondary institutions at the end of the sophomore year. Scores would be compared to SAT entry scores, and the difference between the two exams would be used to judge the general education component of the state's colleges and universities. At this time, preliminary work for the testing program has not been funded.

The Virginia State Council of Higher Education report "Measures of Student Achievement" reviews existing assessment practices at Virginia postsecondary institutions. The Advisory Commission to the Council of Higher Education recommended against a statewide minimum competency testing program for Virginia postsecondary students. Nevertheless, all state postsecondary institutions are to establish campus-designed procedures for measuring student achievement. The state has worked long to develop a diverse set of colleges, the report comments, and the commission did not wish to endanger or threaten that diversity. The Colorado higher education reorganization plan requires each state university and college to measure student achievement. Each institution may adopt its own assessment system, depending on its mission. Schools that are not in compliance by 1990 will lose 20 percent of their appropriation.

At this time the Utah Board of Regents is in the process of revising and updating the Master Plan for Higher Education in Utah. Assessment of student outcomes is under consideration as part of the revision process. Currently, postsecondary institution are strongly encouraged by the Board of Regents to be more active in assessing student learning. Institutions report to the board regarding assessment practices and the improvement of assessment activities. Institutions must submit annual reports of student outcomes for the purposes of program and institutional improvement. Particular assessment instruments or timing of tests are not mandated by the regents. Utah colleges and universities are following a number of assessment strategies. Entry-level testing for advising and placement purposes is used by several schools. One institution has begun to use the ACT COMP for general education testing. Other Utah colleges survey graduating seniors regarding their perception of campus services, coursework, and curricula. In addition, graduates are asked to evaluate the level of skills acquired through their college education.

Entry-level testing for placement and remediation purposes is found at many postsecondary institutions. The New Jersey Board of Higher Education prohibits students from enrolling in college-level courses unless basic skills

proficiencies for the course has been demonstrated through testing; remedial work is noncredit. Students entering Alabama State University are tested for placement purposes, and a "Four Years Plus" program assists seriously underprepared students (Adelman 1984). The City University of New York tests skills of incoming freshmen in reading, writing, and math. While the test is initially given for the purpose of placing low-performing students in remedial classes, it also is used as a gatekeeper for the junior year. Students failing to attain a certain level of proficiency by their junior year may not enroll further.

Ohio has developed a model testing program for mathematics and English placement. The Early Mathematics Placement Program test is administered statewide to high school juniors. Test results let students know their level of readiness for college mathematics while they still have a year left in high school to catch up. The program saves remediation money at the college level and provides motivation for high students students. Remedial math enrollments at Ohio State University are down from 40 percent in 1981 to 25 percent in 1984 (Gov. Richard Celeste). The Early English Composition Assessment Program was first tested in 1984-85. In addition to testing, workshops were offered to high school teachers to improve their teaching effectiveness skills for writing.

Basic skills proficiency, particularly in English, is considered rudimentary for college students. Consequently, a few states and specific institutions test for these skills. The Florida CLAST, Georgia's Regent Exam, and the California Writing Assessment Program have been discussed above. New Jersey is developing a sophomore level test for informational rather than gatekeeping purposes. Exam results should bolster public confidence in higher education and gather objective measures of individual and institutional performance. Both the University of Arizona and the University of Massachusetts at Boston require evidence of writing proficiency midpoint in a student's academic career. Old Dominion University, in Virginia, requires a passing score on the "Exit Writing Exam" for graduation.

## Value-Added Assessment

Value-added assessment has been practiced at Northeast Missouri State University since the early 1970s. It has been a popular and effective program now being considered by other states and universities. The South Dakota Regents of Higher Education has adopted the a similar approach for all state colleges and universities in the system. Kean College, in New Jersey, with a strong minority and nontraditional student population, is instituting a value-added assessment program beginning with the General Education program. Computer-based applications throughout the academic disciplines will accompany the program. The State University of New York at Plattsburgh is in the process of developing a value-added assessment program. The 1985 "Strategic Plan for Higher Education in Kentucky," adopted by the Kentucky Council of Higher Education, included a recommendation for a value-added assessment program, on a pilot basis. The council recommended the assessment process include feedback to Kentucky secondary schools to let them know how their high school graduates fare in college. The legislature has not funded the program at this time. Piedmont Community College in Virginia has a value-added assessment program.

In May 1986, the Texas Committee on Testing of the Coordinating Board of the Texas College and University System voted against adopting a value-added assessment program for Texas public postsecondary institutions. Instead, the committee recommended a basic skills proficiency exam administered to all entering freshmen that must be passed prior to enrollment in upper division coursework. Exam results will not be used to bar community college graduation. The recommendation will be considered by the Coordinating Board and the Select Committee on Higher Education in Summer 1986, and the Texas legislature must grant final approval.

Value-added assessment as practiced at NMSU has undergone several revisions through the years in efforts to constantly upgrade the program. As a result, it is said that assessment permeates the campus in a positive and thoughtful fashion. Although the program was initiated by the administration, faculty play a very active role in it. As presently constituted, the ACT exam is given to half of all incoming freshmen; the ACT COMP (objective test) is given to the other half. At the end of their sophomore year, half the sophomores take the ACT, and the remainder, the COMP. Individual scores are tracked, providing students and their advisors with solid information upon which to recommend further course work for the student. Graduating seniors take nationally normed exams in their major; for example, accounting students take the AICPA Level II, and education majors take the National Teachers Examination. Attitudinal surveys are given to freshmen at entrance, all students biennially, and to seniors upon graduation, supplementing the examinations. Rather than scaring students away, the university has gained a considerable reputation, with increasing numbers of applicants and higher average ACT scores. Assessment at NMSU is used in a positive sense -- no one is penalized for test scores -- to revise the curriculum and strengthen academic offerings. The dialogue of outcome-based education is ongoing at NMSU, affecting student and faculty development, university funding, and long-range planning.

## Student Performance Orientation

Alverno College has a comprehensive assessment program which covers a student's entire college career. The motto of assessment at Alverno is "to serve the learner." Beginning at the freshman level, students are expected ultimately to meet four levels of performance in seven areas of general abilities and to achieve level six in one area of general ability. Students' abilities evaluated through observation and judgment of samples of performance; explicit criteria are established by which to judge performance. Feedback in a central feature of the system so students may learn how effectively the performance is. All of this is "administered sequentially in relation to a learner's development" (Loacker, Cromwell, and O'Brien). Students learn the skill of self-evaluation through this process. Faculty share the responsibility for designing appropriate assessment for the individual classroom; "structured feedback constitute major elements of the assessment design." Alverno students undergo a comprehensive assessment near the end of the sophomore year to determine how students can integrate abilities developed at this stage of college. For example, a half-day assessment might require a student to read material and present a synthesis of

the material orally; to work with a desk top of letters, phone messages, and memos; and to work with four other students to develop guidelines and recommendations, based on a clear rationale (Loacker, Cromwell, and O'Brien). Assessment in the major field are performance oriented. External assessors from the community assist in the evaluation process. Alverno students may go through more than 100 performance evaluations throughout college (Harris).

Other examples of comprehensive assessment programs are found in professional schools. For example, the University of Montana Law School has revised its curriculum to a more pragmatic, performance-oriented program. Students are associated with simulated "law firms" and evaluation and feedback are provided throughout the program. The William Mitchell College of Law in Minnesota emphasizes classroom-level assessment, practical skills courses, and student representation of a real client with faculty supervision. At Harvard, the Oliver Wendell Holmes Society New Pathway Project has been initiated in the General Medical Education program (Elman and Lynton). A randomly selected group of students are put through a competency-based program that included mastery of knowledge in addition to competency with programmed patients. Students will be compared with those in the more traditional program.

## Performance-Based Funding

Performance-based funding has been adopted in Tennessee on a statewide basis since 1979 in order to stimulate public postsecondary institutions to meet five performance criteria, including the assessment of student outcomes. In the early years of the program, state colleges and universities could receive a bonus of 2 percent of nonearmarked funding for meeting the standards. For an institution such as Tennessee Tech, the program netted an additional $200,000 to $300,000 per year ("Let's Reward Quality"). Beginning in 1984, incentive money was increased to 5 percent. Student outcomes criteria are measured by student performance on general education measures, subject matter mastery, and indicators of student satisfaction with their college (Ewell, "Levers for Change"). Some institutions use the ACT COMP to assess the effectiveness of general education and GRE subject matter or locally developed major exams to evaluate subject matter mastery.

Tennessee has enacted other postsecondary improvement programs serving as models for state action. Challenge grants for endowed chairs have been funded. In addition, the "Bragg Marks" program encourages Tennessee institutions to establish goals on fourteen measures, such as first attempt pass rates on licensing exams and student retention rates. By 1990 each public college and university is expected to demonstrate improvement on the fourteen criteria.

Other states have initiated varieties of performance funding programs. Missouri has reward individual institutional improvement based on improved student performance. New York enacted an incentive grant program (Bundy Funds) for private institutions based on the number of degrees conferred (Ewell, "Levers for Change). Massachusetts provides funds for Centers of Excellence. Ohio has adopted five specific grant programs to improve higher education, including funding Centers for Excellence and chaired

professorships. While these programs are not specifically based on student outcomes, they provide incentives for institutional improvement. New Jersey also has provided special funding to institutions; the Governor's Challenge Funds target state institutions demonstrating dramatic improvement. South Dakota has started a performance funding program, and Florida has adopted an initiative to fund Indicators of Progress toward Excellence in Education (Pickens). The Virginia Fund for Excellence provides grants for institutionally generated, goal-oriented projects. For example, Virginia Commonwealth University received $120,000 to implement a program for student writing skill improvement. James Madison University has received a grant of $125,000 to begin "Initiatives for Excellence and Accountability: A Five Year Plan." The plan includes a comprehensive review of curricula. A university wide evaluation and assessment system to measure student achievement is being explored (Report of Study Conducted by the Council for Higher Education of Virginia).

State higher education audit and program review procedures are other mechanisms a state can use to encourage the assessment of student outcomes in judging institutional and departmental effectiveness. North Carolina requires campuses to submit data on student progress. The Hawaii community college program review process requires the college to submit data on student placement and performance. Maryland and Oregon have adopted similar programs. Florida and New Jersey have unit record systems allowing for uniform monitoring of students through the public university system.

## Other Examples of Assessment Programs

The State University of New York (SUNY) at Albany requires each department to provide data on alumni success and alumni perceptions of the college experience as part of the regular budget cycle process (Ewell, "Levers for Change"). SUNY - Potsdam piloted an assessment program in 1981 to discover student skills in five areas: humanities, fine arts, social science, natural science, and mathematics. Using students from three divisions of the Universities, ETS's measures for general education, supplemented by local assessment techniques, were administered. For example, in the drama area, The Birdbath by Leonard Melfi was presented, and students were asked to respond to the drama's value structure. An additional aspect of the university's assessment program was the creation of institution-specific instruments applicable in virtually every classroom. The assessment program, part of the Academic Program Evaluation Project of American Association of State Colleges and Universities has lain dormant for several years. However, the faculty is in the process of reviving assessment as part of a general education curricular revision emphasizing outcomes. The United States Military Academy at West Point utilizes comprehensive assessment to determine how effectively the academy is preparing graduates for military leadership roles. The emphasis is on leadership skills rather than educational progress. Multiple measures assess student leadership development, accompanied by a feedback process.

## CONCLUSION

Evaluating the quality of a college education solely by counting books in the library or microscopes in the biology lab is rapidly becoming outmoded.

The major outcome of postsecondary education should be clear: student learning and performance. Policymakers, consumers, state higher education authorities, students, and parents have too high an investment in the higher education system, public and private, to continue to rely on superficial and synthetic measures of student performance. Model assessment programs have been developed by a handful of colleges and universities and states. Now other states and postsecondary institutions are following suit, adapting programs and policies to their own missions and higher education systems. Measurement instruments are being developed locally and nationally to supplement the already-existing stock. At last, assessment of student learning on a systematic basis has been added to the agenda of postsecondary education and the states.

REFERENCES

Abernathy, L. J. "Institutionally Based Student Assessment for College Placement." The College Board Review, No. 138, Winter 1985-86, pp. 9-11, 29.

Adelman, Clifford (ed.). "From Reports to Response: Proceedings of Regional Conferences on the Quality of American Higher Education," Washington, D.C.: U.S. Department of Education, 1986.

Adelman, Clifford (ed.). "Assessment in American Higher Education: Issues and Context." Washington, D.C.: U.S. Department of Education, Office of Educational Research Improvement, 1986.

Adelman, Clifford. Senior Associate, U.S. Department of Education, Office of Education Research Improvement. Interview, March 28, 1986.

Adelman, Clifford. "The Standardized Test Scores of College Graduates 1964-1982." Prepared for the Study Group on the Condition of Excellence in American Higher Education. Washington, D.C.: U.S. Department of Education, National Institute of Education, 1984.

Adelman, Clifford. "Starting With Students: Promising Approaches in American Higher Education." Presented to the Study Group on the Conditions of Excellence in American Higher Education. Washington, D.C.: U.S. Department of Education, National Institute of Education, 1985.

American Association of State Colleges and Universities. Statement to the National Governors' Association Task Force on College Quality, 1985.

The American College Personnel Association. "Affirming Individual Differences in the Learning Process: A Model for Improving College Quality." A paper submitted to the National Governors' Association Task Force on College Quality, 1986.

Anrig, Gregory R. "A Message for Governors and State Legislators: 'The Minimum Competency Approach Can Be Bad for the Health of Higher Education.'" State education Leader 5, forthcoming, September 1986.

Ashcroft, John, Governor of Missouri. "How Can You Measure College Learning?" Governors' Weekly Bulletin, Vol. 26, September 20, 1985, p. 3.

Association of American Colleges. Integrity in the College Curriculum: A Report to the Academic Community: The Findings and Recommendations of the Project on Redefining the Meaning and Purpose of Baccalaureate Degrees. Washington, D.C.: Association of American Colleges, 1985.

Astin, Alexander, and Paros, Robert. The Educational and Vocational Development of College Students. Washington, D.C.: American Council on Education, 1969.

Bailey, Adrienne. "College Readiness from the Viewpoint of the Educational Equality Project." Testimony to the National Governors' Association Task Force on School Readiness, December 31, 1985.

Banta, Trudy W. "Use of Outcomes Information at the University of Tennessee." In Assessing Educational Outcomes, Peter T. Ewell (ed.), New Directions for Institutional Research, No. 47. San Francisco, Calif.: Jossey-Bass, Inc., 1985.

Bernstein, Melvin H., and the New England Board of Higher Education. "Higher Education and the State: New Linkages for Economic Development." Washington, D.C.: National Institute for Work and Learning, 1986.

Bogue, E. Grady, and Brown, Wayne. "Performance Incentives for State Colleges." Harvard Business Review, Vol. 60, November/December 1982, pp. 123-128.

Bowen, Howard R., and Schuster, Jack H. American Professors: A National Resource Imperiled. New York, N.Y.: Oxford University Press, 1985.

Boyer, Carol M., and McGuiness, Aims C. Jr. "Recent State Initiatives to Improve the Quality of Undergraduate Education." State education Leader, 5, Winter 1986, p. 12.

Carnes, Bruce M. Deputy Undersecretary of Education, U.S. Department of Education. Testimony to the National Governors' Association Task Force on College Quality. Washington, D.C., 1986.

Celeste, Richard F., Governor of Ohio. Paper submitted to the National Governors' Association Task Force on College Quality, 1986.

The College Board. "Academic Preparation for College: What Students Need to Know and Be Able to Do." New York, N.Y.: The College Board, 1983.

"Colleges Test Their Work." New York Times, November 5, 1985, pp. C1, C7.

The Council on Postsecondary Accreditation. "Educational Quality and Accreditation: A Call for Diversity, Continuity and Innovation." Washington, D.C.: The Council on Postsecondary Accreditation, 1986.

Davila, Evelyn M. Project Director. "Today's Urban University Students: Part I, Profile of a New Generation." New York, N.Y.: College Board, 1985.

Diamond, Robert M., and Gray, Peter J. "A Practical Approach to Course and Curriculum Improvement." Syracuse, N.Y.: Center for Instructional Development, Syracuse University, n.d.

Doyle, Arthur J. "College Entry Level Skills Assessment." Paper presented to the National Governors' Association Task Force on College Quality, 1985.

Doyle, Arthur J. "Computerized Placement Tests." Testimony to the National Governors' Association Task Force on Educational Technology, 1985.

Doyle, Arthur J. "Computerized Diagnostic Testing." Testimony to the National Governors' Association Task Force on Educational Technology, 1985.

Education Commission of the States. "Catalog of Changes: Incentives for Quality and Management Flexility in Higher Education." Denver, Colo.: Education Commission of the States, 1984.

Education Commission of the States. "State Initiatives: An Update." Change, Vol. 17, November/December 1985, pp. 16-17.

Edgerton, Russell, and Marchese, Theodore. "Leadership for Assessment." Unpublished paper. Washington, D.C.: American Association for Higher Education, 1986.

Ehrmann, Stephen C. "Teaching with Tests: Why, and Why Now?" Washington, D.C.: Fund for the Improvement of Postsecondary Education, 1984.

El-Khawas, Elaine. "Campus Trends, 1984." Higher Education Panel Report No. 65. Washington, D.C.: American Council on Education, 1985.

El-Khawas, Elaine. "Campus Trends, 1986." Higher Education Panel Report No. 73. Washington, D.C.: American Council on Education, 1986.

Elman, Sandra E., and Lynton, Ernest A. "Assessment in Career-Oriented Education." In Assessment in American Higher Education: Issues and Contexts. Clifford Adelman (ed.). Washington, D.C.: U.S. Department of Education, 1986.

Elofson, Stan. "Colorado Reorganizes Higher Education." State education Leader, 4, Fall 1985, pp. 4-5.

ERIC Digest. "Educational Accountability." Los Angeles, Calif.: ERIC Clearinghouse for Junior Colleges, 1984.

Estes, David R. College Relations Manager, McDonnell Douglas Corporation. Testimony to the National Governors' Association Task Force on College Quality. Washington, D.C., 1986.

Ewell, Peter. "Assessment: What's It All About?" Change, 17, November/December 1985, pp. 32-36.

Ewell, Peter T. "Four Myths About Assessment." Testimony to the National Governors' Association Task Force on College Quality. Kansas City, Mo., 1986.

Ewell, Peter T. "The Legislative Role in Improving Postsecondary Education." Education and Job Training Series, State Legislative Report, Vol. 10, No. 14. Denver, Colo.: National Conference on State Legislatures, 1985.

Ewell, Peter T. "Levers for Change: The Role of State Government in Improving the Quality of Postsecondary Education." Denver, Colo.: Education Commission of the States, 1985.

Ewell, Peter T. "Some Implications for Practice." In Assessing Educational Outcomes, Peter T. Ewell (ed.) New Directions for Institutional Research No. 47. San Francisco, Calif.: Jossey-Bass, Inc., 1985.

Folger, John (ed.) Financial Incentives for Academic Quality. New Directions for Institutional Research No. 48. San Francisco, Calif.: Jossey-Bass, Inc., 1984.

"Freshmen Face Stiffer Requirements." Ohio State Lantern, June 4, 1985, n.p.

Goodwin, Fred B. Testimony to the National Governors' Association Task Force on College Quality. Kansas City, Mo., 1986.

Hanford, George H. "The SAT and Statewide Assessment." Remarks delivered to the Eighth Northern New England Educational Tests, Measurement and Evaluation Conference. Plymouth, N.H., March 26, 1986.

Hanford, George H. Testimony to the National Governors' Association Task Force on College Quality. Washington, D.C., 1986.

Harris, John. "Assessing Outcomes in Higher Education." In Assessment in American Higher Education: Issues and Contexts. Clifford Adelman (ed.). Washington, D.C.: U.S. Department of Education, 1986.

Harris, John. Professor, David Lipscomb College. Interview, April 11, 1986.

Hartle, Terry W. "The Growing Interest in Measuring the Educational Achievement of College Students." In Assessment in American Higher Education: Issues and Contexts. Clifford Adelman (ed.). Washington, D.C.: U.S. Department of Education, 1986.

"High School Graduates: Projections for the Fifty States." Western Interstate Commission for Higher Education, 1984.

Hodgkinson, Harold L. "All One System: Demographics of Education, Kindergarten through Graduate School." Washington, D.C.: Institute for Educational Leadership, Inc., 1985.

Hollander, Chancellor T. Edward. "College Outcomes Evaluation Program." Memorandum to New Jersey Board of Higher Education Academic Affairs Committee, May 17, 1985.

Jacobson, Robert L. "Most Students Are Satisfied with Their Education, Survey Indicates, but Frustrations Are Widespread." Chronicle of Higher Education, Vol. 31, February 5, 1986, pp. 1, 30-31.

Jacobson, Robert L. "New Carnegie Data Show Faculty Members Uneasy About the State of Academe and Their Own Careers." Chronicle of Higher Education, Vol. 31, December 18, 1985, pp. 1, 24, 26-28.

Jaschik, Scott. "As States Weigh 'Value-Added' Tests, Northeast Missouri Offers Model." Chronicle of Higher Education, October 2, 1985.

Kean, Thomas, Governor of New Jersey. Interview by Frank Newman. "Rising Expectations: Can States Help Review Quality?" Change, Vol. 17, November/December 1985, pp. 12-17.

Kelly, J. Terrence. Testimony to the National Governors' Association Task Force on College Quality, St. Louis, Mo., 1985.

Kinnick, Mary K. "Increasing the Use of Student Outcomes Information." In Assessing Educational Outcomes. Peter T. Ewell (ed.). New Directions for Institutional Research No. 47. San Francisco, Calif.: Jossey-Bass, Inc., 1985.

Koretz, Daniel. "Trends in Educational Achievement." Washington, D.C.: The Congress of the United States, Congressional Budget Office, 1986.

Kosloff, Jessica. "Survey of Recent State Initiatives to Improve the Quality of Undergraduate Education." A briefing paper prepared for the Education Commission of the States, 1985.

Lawrence, Judith K., and Green, Kenneth C. A Question of Quality: The Higher Education Ratings Game. ERIC/Association for Higher Education Report 5. Washington, D.C.: American Association for Higher Education, 1980.

Learning, Assessment, Retention Consortium of California. "California Community Colleges as Providers of Remediation." Field draft. Sacramento, Calif.: California Association of Community Colleges, 1985.

"Let's Reward Quality: Tennessee's Bold Experiment." Change, Vol. 17, November/December 1985, pp. 37-45.

Levy, Robert A. "Development of Performance Funding Criteria by the Tennessee Higher Education Commission: A Chronology and Evaluation." In Performance Funding in Higher Education: A Critical Analysis of Tennessee's Experience. Trudy W. Banta (ed.) with an assessment by Peter Ewell. Boulder, Colo.: National Center for Higher Education Management Systems, 1986.

Loacker, Georgine; Cromwell, Lucy; and O'Brien, Kathleen. "Assessment in Higher Education: 'To Serve the Learner.'" In Assessment in American Higher Education: Issues and Contexts. Clifford Adelman (ed.). Washington, D.C.: U.S. Department of Education, 1986.

Losak, John. Testimony to the National Governors' Association Task Force on College Quality. St. Louis, Mo., 1985.

Low, Stephen. Testimony to the National Governors' Association on College Quality, Washington, D.C., 1986.

Magrath, Peter. Testimony to the National Governors' Association Task Force on College Quality. St. Louis, Mo., 1985.

Marchese, Ted. "Learning About Assessment." AAHE Bulletin, 38, September 1985, pp. 10-13.

Marchese, Ted. Speech to the Learning, Assessment, Retention Consortium of California. March 10, 1986.

Marcus, Laurence R.; Leone, Anita O.; and Goldberg, Edward D. Quality Assurance in Higher Education. ASHE-ERIC Higher Education Research Report. Washington, D.C.: Association for the Study of Higher Education, 1983.

McCabe, Robert H. "Legacy of Shame." Miamia/South Florida Magazine, November 1985, p. 82.

McClain, Charles J. Testimony to the National Governors' Association Task Force on College Quality. St. Louis, Mo., 1985.

McClain, Charles J. and Krueger, Darrell W. "Using Outcomes Assessment: A Case Study in Institutional Change." In Assessing Educational Outcomes. Peter T. Ewell (ed.) New Directions for Institutional Research No. 47. San Francisco, Calif.: Jossey-Bass, Inc., 1985.

McGuinness, Aims, Jr. "Action Can Improve Undergraduate Quality." State education Leader, 4, Summer 1985, pp. 1-2.

McNett, Ian. "Demographic Imperatives: Implications for Educational Policy." Report of the June 8, 1983, forum on "The Demographics of Changing Ethnic Populations and their Implications for Elementary-Secondary and Postsecondary Educational Policy," sponsored by the American Council on Education, Forum of Educational Organization Leaders, and Institute for Educational Leadership.

Maryland State Board for Higher Education. "Alternatives for Testing Maryland College Students." Annapolis, Md.: Maryland State Board for Higher Education, 1984.

Mentkowski, Marcia, and Loacker, Georgine. "Assessing and Validating the Outcomes of College." In Assessing Educational Outcomes. Peter T. Ewell (ed.) New Directions for Institutional Research No. 47. San Francisco, Calif.: Jossey-Bass, Inc., 1985.

Millard, Richard. Testimony to the National Governors' Association Task Force on College Quality. Washington, D.C., 1986.

Mingle, James R. "Measuring the Educational Achievement of Undergraduates: State and National Developments." Denver, Colo.: State Higher Education Executive Officers, 1985.

National Center for Education Statistics. The Condition of Education, 1984 Edition. Washington, D.C.: U.S. Department of Education, 1984.

National Commission on Excellence in Education. A Nation at Risk: The Imperative for Educational Reform. Washington, D.C.: U.S. Government Printing Office, 1983

National Commission on Higher Education Issues. To Strengthen Quality in Higher Education: Summary Recommendations of the National Commission on Higher Education Issues. Washington, D.C.: National Commission on Higher Education Issues, 1982.

National Endowment for the Humanities (NEH). To Reclaim a Legacy. A report written by William J. Bennett, chairman, NEH. Based on findings of the Study Group on the State of Learning in the Humanities in Higher Education. Washington, D.C.: National Endowment for the Humanities, 1984.

National Institute of Education. Involvement in Learning: Realizing the Potential of American Higher Education. Study Group on the Conditions of Excellence in American Higher Education. Washington, D.C.: National Institute of Education, 1984.

New Jersey Basic Skills Council. "Effectiveness of Remedial Programs in New Jersey Public College and Universities Fall 1982-Spring 1984." Trenton, N.J.: New Jersey Department of Higher Education, 1985.

New Jersey Basic Skills Council. "Report on the Character of Remedial Programs in New Jersey Public Colleges and Universities." Trenton, N.J.: New Jersey Department of Higher Education, 1985.

Newman, Frank. "States Join the Debate." Change, November/December 1985, pp. 6-7.

Nielsen, Robert. Testimony to the National Governors' Association Task Force on College Quality. Kansas City, Mo., 1986.

Northeast Missouri State University. "The Value Added Program: Assessing Undergraduate Education." Prepared for the National Governors' Association Task Force on College Quality. St. Louis, Mo., 1985.

Novak, Richard. "Efforts to Restructure State Higher Education Governance Continue." Unpublished paper, American Association of State Colleges and Universities, 1985.

Office of Minority Concerns. "Fourth Annual Status Report of Minorities in Higher Education." Washington, D.C.: American Council on Education, 1985.

Ohio Board of Regents. "Ohio Steps Out Front...An Investment in Excellence." Columbus, Ohio: Ohio Board of Regents, 1985.

Office of Academic Affairs. "The City University of New York Freshman Skills Assessment Program: Student Information Bulletin." New York, N.Y.: The City University of New York, n.d

Pickens, William H. "State Budgeting for Higher Education: Is There an Ideal Approach?" The Grapevine. Normal, Ill: Center for Higher Education, Illinois State University, January/February 1986, pp. 2041-2045.

Preer, Jean L. Competence, Admissions, and Articulation: Returning to the Basics in Higher Education. ASHE-ERIC Higher Education Research Report No. 6. Washington, D.C.: Association for the Study of Higher Education, 1983.

Prince, Colonel Howard. Testimony to the National Governors' Association Task Force on College Quality. Washington, D.C., 1986.

Proceedings of the 1984 Educational Testing Service Invitational Conference. Educational Standards, Testing and Access. Princeton, N.J.: Educational Testing Service, 1985, pp. 15-22.

Read, Joel Sr. President, Alverno College. Testimony to the National Governors' Association Task Force on College Quality. St. Louis, Mo., 1985.

Report of Study Conducted by the Council of Higher Education for Virginia. "The Measurement of Student Achievement and the Assurance of Quality in Virginia Higher Education." Senate Document No. 14. Richmond, Va.: Commonwealth of Virginia, 1986.

Smith, Hoke. President, Towson State University. Testimony presented to the National Governors' Association Task Force on College Quality. Washington, D.C., 1986.

The South Dakota Board of Regents. "The South Dakota Board of Regents Assessment Program." Mimeographed, n.p.: n.p., 1986.

Southern Regional Education Board. "Access to Quality Undergraduate Education." A Report to the Southern Regional Education Board by its Commission for Educational Quality. Atlanta, Ga.: Southern Regional Education Board, 1985.

Southern Regional Education Board. "Access to Quality Undergraduate Education." A Report to the Southern Regional Education Board by its Commission for Educational Quality. Atlanta, Ga.: Southern Regional Education Board, 1985.

"State Initiatives: An Update." Change, 17, November/December 1985, pp. 16-17.

Task Force on Education for Economic Growth. Action in the States: Progress toward Education Renewal. Denver, Colo.: Education Commission of the States, 1984.

Taylor, Terry. "A Value-Added Student Assessment Model: Northeast Missouri State University." Economics of Education Review, 4, 1985, pp. 341-350.

Tennessee Higher Education Commission. "Allocation of State Funds on a Performance Criterion." The Report of the Performance Funding Project. E. Grady Bogue, Project Director. Nashville, Tenn.: Tennessee Higher Education Commission, 1980.

"Toward Educational Excellence." Report of the Task Force on Initiatives Toward Educational Excellence. Kansas City, Mo.: The Metropolitan Community College District, 1985.

Turnbull, William W. "Are They Learning Anything in College? Change, 17, November/December 1985, pp. 23-26.

U.S. Census Population Reports. Series P20 No. 404. Bureau of the Census, U.S. Department of Commerce. Washington, D.C., 1980.

U.S. Census Population Reports. Series P25 No. 952. Bureau of the Census, U.S. Department of Commerce. Washington, D.C., 1980.

University of Montana School of Law. "Academic Planning Project: Interim Report." Billings, Mont.: University of Montana School of Law, n.d.

Vobejda, Barbara. "State Test Debated to Measure Colleges." The Washington Post, October 17, 1985.

Warren, Jonathan R. (ed.). Meeting the New Demand for Standards. San Francisco, Calif.: Jossey-Bass, Inc., 1983.

Young, Candy Cartwright. "Faculty and Educational Assessment." Testimony to the National Governors' Association Task Force on College Quality. Kansas City, Mo., 1985.

# THE STATE ROLE IN ASSESSING COLLEGE OUTCOMES: POLICY CHOICES AND PROBABLE IMPACTS

Prepared by Peter T. Ewell
Senior Associate
National Center for Higher Education Management Systems

Assessing the outcomes of college and university instruction is a topic at the center of interest and debate in most discussions of state policy on improving undergraduate education. Proponents claim that assessment will revitalize the classroom, will allow continuous and positive improvement of curriculum, and will provide a more appropriate focus for external institutional accountability. Opponents claim that accurate assessment of a process so complex as college-level instruction is impossible, and that narrowly conceived attempts to implement assessment will do irreparable damage to academic vitality. Neither position has much evidence available to support it. The history of explicit, state-mandated assessment programs is brief: to date, only seven states have mandated assessment as part of higher education reform legislation or policy, and only two of these efforts have been in place for more than a year. As a result, no state has acquired sufficient experience with the concept to satisfactorily inform current practice.

From the outset, however, it is important that state policymakers recognize that "assessment" has many meanings in current discourse. Consequently, developing a statewide policy on assessment entails making a series of concrete choices about what to assess and how to conduct such a program. As public policy alternatives, for example, it is difficult to imagine approaches more distinct than those reflected in the two most established statewide programs -- Florida and Tennessee. The former emphasizes uniform testing of individual students in basic areas of minimum competency at different points in their academic careers. The latter emphasizes institutional accountability by providing incentives for colleges and universities to collect, report and use a range of institution-specific data on instructional quality. Informed public policy demands that the rationale underlying such differences be made clear, and that the probable consequences of different approaches be fully spelled out.

Given these demands, the purpose of this paper is to present available evidence about the impacts of different state-mandated approaches. Because of the relative scarcity of established state programs, however, evidence will be drawn primarily from individual institutions that have been experimenting with assessment sufficiently long for the consequences of their efforts to develop. We at the National Center for Higher Education Management Systems have worked with some two dozen such institutions and can generalize effectively from this base of experience (see Appendix A). Further, much can be learned about the potential impact of state- sponsored assessment by carefully examining the consequences of other, similar reforms. Such efforts include statewide program review and approval processes, statewide cost studies, and special accreditation of professional programs.

The structure of the argument will be as follows. First, some of the most common claims for and against assessment will be listed and examined. A primary purpose of this brief exposition will be to separate the rhetoric of these claims from the underlying reality of probable impact on programs, institutions, and individual student learning. Two of the most important claims against assessment, which we will explicitly revisit in the light of available evidence, are that explicit testing programs will excessively narrow college curriculum offerings and will induce college faculty to "teach to the test." Secondly, some current state approaches to assessment will be

examined, and each approach will be dissected in terms of a set of explicit policy choices that need to be made. Finally, available evidence on the probable impact of different approaches will be assembled and discussed in three critical areas: the curriculum, the individual classroom, and the effective functioning of the college or university as a self-governing organization.

In conclusion, the paper will offer a set of summary policy recommendations intended to structure discussion in particular state settings. A major theme is the need for states to undertake comprehensive reform efforts -- efforts of which assessment is an integral part -- rather than mandating assessment in isolation of other new initiatives. A second theme is the need for states to carefully balance their use of authority with the inherent limits of accountability. What is needed are policies that induce change, but that do not at the same time create levels of institutional resistance that prevent meaningful local action. Higher education institutions have traditionally been self-governing entities and the academic enterprise a self-policing one. Because organizational structures tend to be decentralized, they can be induced to change by external incentives that promise improved effectiveness, but that do not simultaneously threaten traditional decisionmaking autonomy. Achieving such a balance in its approach to assessment is perhaps the most challenging task of any state.

## THE ASSESSMENT DEBATE: PROS AND CONS REVISITED

One of the primary obstacles to a dispassionate discussion of the probable impacts of assessment is the level of rhetoric associated with positions for and against it. This is partly because of a strong diversity of motive among those who have made such claims. It is partly because both proponents and opponents, as is true in any debate, tend to base their respective arguments on different concepts of what assessment will actually look like. Proponents generally view assessment in highly idealized terms and ignore many of the practical difficulties in implementing such programs effectively and fairly. Opponents, in contrast, emphasize such difficulties and see assessment in terms of "worst possible case" scenarios of how things might go wrong. Neither view can make much progress until its premises are clarified. Once this is accomplished, it is possible to array existing evidence about the impact of assessment to test the degree to which both positive and negative claims can be sustained.

### 1. Claims for Assessment

Despite a wide diversity of sources, claims about the benefits of assessment tend to cluster around four basic positions:

- o Assessment programs are important in providing "quality assurance" about higher education's effectiveness. From this perspective, the primary beneficiaries of higher education are the "consumers" of its products -- for example, potential students and their parents, or employers interested in the actual skills possessed by college graduates. A second beneficiary is the public at large (and its elected representatives) who are vitally interested in judging the

"return on investment" of an increasingly diverse and expensive public enterprise. Examples of such positions range from a recent statement by the Secretary of Education that colleges should measure outcomes to better inform potential student choices, to the claims of individual institutions to be providing "degrees with integrity" (AASCU 1984).

- o Assessment is a productive "lever for change" in inducing colleges and universities to make improvements consistent with statewide concerns. In this case, the primary target of assessment is the individual college or university, and its response to statewide initiatives. For example, in Tennessee, the five dimensions of institutional performance established by a statewide Instructional Evaluation Schedule serve as tangible reminders of the particular types of performance deemed important by the state (Bogue and Brown 1982). Similarly, Florida's College Level Academic Skills Program (CLASP) is in part designed to unambiguously signal to institutions the kinds of skills considered essential.

- o Assessment enables colleges and universities to identify their individual strengths and weaknesses, and to make changes consistent with their own goals and purposes. The underlying argument here is that the best academic programs evolve through a continuous, institution-specific process of self-examination and renewal. Examples of this position can be drawn from each of the recent national reports on improving undergraduate instruction, for example, the American Association of Colleges' "Integrity in the College Curriculum" (1985). Others have seen this "self-regarding" approach as a key to maintaining institutional quality while preserving needed diversity among institutional types (Ewell 1984).

- o Assessment enables individual students to learn more and learn more quickly by providing periodic, concrete feedback on their performance. The primary beneficiary of assessment here is held to be the individual student. The assumed mechanism behind this position -- that individual learners benefit from explicit, performance-based feedback provided by an integral part of instruction by faculty or by specially trained assessors -- is backed by considerable educational laboratory and field research at the elementary and secondary levels. In current national discourse, this position is exemplified by NIE's "Involvement in Learning" (1984), which calls for all colleges and universities to assess student performance independent of the grading process, and by Alverno College, an institution that adopted a comprehensive competency-based curriculum founded on these principles a decade ago.

## 2. Objections to Assessment

Common arguments against assessment, particularly as an explicit element of state policy, are in contrast much more diffuse. Most objections, however, can be grouped under one of the following:

- Assessment will excessively narrow the college curriculum by discouraging institutions to offer (and faculty to teach) courses whose intended outcomes or instructional methods are not easily susceptible to quantitative evaluation. The essence of this position is that the very act of prescribing a set of outcomes dimensions inherently devalues an experience that should be holistic and integrative. Its logic is similar to discussions of the negative impact of specialized accreditation, a process that increasingly demands that curricula consist of courses leading to particular professional skills (for example, NIE 1984, Bennett 1984). By specifying only a narrow range of outcomes, courses that stress the development of more general and difficult to assess abilities will be crowded out of the curriculum. Even if there is no official mechanism for this, student demand will be greatest in those courses deemed most beneficial in terms of test performance. As a result, opponents fear, courses not directly related to test performance will be undercontrolled and may eventually for this reason be curtailed.

- Assessment limits the kinds of learning that will take place in college classrooms by encouraging instructors to "teach to the test." In essence, this objection holds that assessment will change the nature of classroom instruction by limiting the range of methods and choice of subject matters provided by individual teachers. Consequently, such objections are often clothed with the traditional rhetoric of "academic freedom" -- a principle at the foundation of the academic structure. Conceived originally as an extension of First Amendment rights, this principle holds that the individual faculty member, as a professional, should be allowed full autonomy in the classroom. By limiting the professional ability of the individual faculty member to exclusively judge the course content and instructional method best suited to communicating a complex subject matter, assessment will encourage students to master only those elements of knowledge that they need to pass a test (NEA 1986).

- By violating important and historic principles of institutional autonomy and decentralization, assessment will limit the ability of colleges and universities to innovate, and to remain on the cutting edge of scholarship and instruction. In essence, this position maintains that holding institutions accountable for a narrowly specified range of results will impair their effective functioning as organizations. This position is distinct from "academic freedom" arguments in that it is the institution rather than the individual that is claimed to have particular rights and obligations that exempt it from the traditional structure of public accountability (Leslie 1986). Support for this position is provided by research on college and university effectiveness that maintains that goal clarity and substantial differentiation of an institution's goals from those of its competitors leads to more effective functioning (for example, Chaffee 1984). Such positions are exemplified by those who fear that violations of autonomy will reduce important differences among institutions (Smith 1986).

- o Assessment is technically infeasible and excessively costly, and diverts institutional attention from other more important activities. This set of "technical" objections holds that any theoretical benefits of assessment are not worth their cost. Some such positions stress the fact that most attempts to implement assessment rely on standardized cognitive achievement tests that were designed for quite different purposes. Others see beneficial assessment as already taking place on most college campuses, and thus see investments in large-scale assessment as a waste of funds. Finally, assessment is seen as adding burdens to an already hard-pressed and underpaid faculty. Such "technical" objections, it is important to note, are rarely advanced on their own merits, and are generally combined with one or more of the objections above.

Differences among these many presumed benefits and costs of assessment are primarily distinctions about who sponsors the process, who controls it, and how it is conducted. Among claimed benefits, for example, all positions maintain that assessment will ultimately improve the learning experiences of individual students, but only the last sees this as a direct effect. The others see this benefit as a result of improved institutional programs and services. Moreover, the first two claimed benefits depend upon external state authority to induce institutions to make needed changes. The last two emphasize the independent role of institutions in identifying and addressing areas of instructional strength and weakness.

Turning to objections to assessment, all four positions largely assume that assessment will be imposed upon institutions from without -- either through state authority or through mechanisms such as regional or specialized accreditation. At the same time, all but technical objections assume that assessment methods will rely heavily on standardized cognitive tests, administered commonly to all the institutions in a state system. Technical objections, generally advanced in conjunction with one or more of the others, reinforce this position by claiming that reasonable methodological alternatives to standardized testing are either infeasible or are excessively costly.

Divergent assumptions about what assessment programs will actually look like highlight the need for a careful review of state policy alternatives. In general, such alternatives can be defined in terms of how each state chooses to answer four basic questions in defining its approach: (1) what is the basic purpose of assessment, (2) who is to be assessed, (3) what methods are to be used in assessment, and (4) who controls and operates the assessment process? At the same time, divergent views of the impact of assessment depend upon where the primary positive and negative impacts are expected -- on the curriculum, in the classroom, or in terms of effective institutional functioning. Each of these issues will be treated more systematically in the sections that follow.

## ASSESSMENT AS PUBLIC POLICY: THREE BASIC APPROACHES

Because the range of state roles in assessment is considerable, it is difficult to make uniform generalizations about probable impact. Moreover, as

noted above, the claims of advocates and opponents of state-mandated assessment generally are based on quite different views of how such programs will be designed and implemented. Proponents have tended to see assessment largely as an institution-specific process, perhaps stimulated by outside authority or incentive, but practiced primarily by institutions themselves. Opponents, in contrast, have raised a spectre of universal state-mandated standardized testing. Few actual programs resemble either extreme. With the exception of a handful of pioneers (most notably Alverno College and Northeast Missouri State University), emerging institutional assessment programs have indeed been designed in reaction to existing or anticipated accountability or accreditation requirements. Most, however, employ a range of assessment techniques, and examine quite diverse dimensions of student and curricular performance. None, to date, has approximated the "worst case" model posited by assessment's opponents.

In practice, of course, all established and emerging state approaches to assessment involve making choices on a set of four dimensions:

- o What is the purpose of assessment? This dimension, in essence, concerns the degree to which the state views assessment primarily as a device to ensure compliance with minimum standards, as a means to achieve broadly stated state objectives, or as a mechanism for inducing ongoing, institution-specific efforts at self-improvement.

- o At whom is assessment directed? This dimension concerns the appropriate "unit of analysis" for assessment. While all assessment depends upon measuring the performance of individual students, is this information to be used to evaluate what individual students know and can do, to identify the strengths and weaknesses of curricula and programs, or to assess the overall instructional effectiveness of entire institutions?

- o How is assessment conducted? This dimension concerns the actual methods to be used to assess educational outcomes. Such methodological issues are generally of three kinds: the degree to which standardized cognitive achievement tests will be used, whether single or multiple sources of assessment evidence will be employed, and the degree to which assessment depends upon "one-shot" testing at the conclusion of a program or involves longitudinal evaluation of student progress from the beginning to the end of a program of study.

- o Who controls assessment? The essence of this choice is the degree to which assessment is mandated, designed, and administered by a central, statewide authority. Areas of concern within this dimension include determination of who is responsible for establishing the basic areas of performance to be assessed, establishing the methods to be used, and actually conducting the program on a day-to-day basis.

While choices on each of these dimensions are in principle independent, emerging experience at the state level suggests three distinct approaches. One path sees state assessment primarily as an act of quality assurance, and therefore tends to concentrate on student-level testing, standardized

procedures employed in common by all institutions, and a relatively narrow range of assessment dimensions -- generally restricted to the attainment of basic collegiate skills. This approach, which in many ways resembles the occupational/professional certification process, is most clearly exemplified by Florida, which requires all students to pass basic skills tests as a condition for advancement, or by Georgia and California, which require minimum skills in such areas as writing to be demonstrated by graduation. By mandating and testing for minimum levels of student performance as a requirement for promotion to junior status, the Florida program is clearly intended as a quality control device. Indeed, it can be seen in the same light as the mandatory skills examinations employed by Florida and other states for screening college applicants. Certainly the CLASP testing program will yield valuable institution and program specific information for future lower division improvement in Florida. Indeed, there is good evidence that a similar diagnostic testing program in Ohio -- targeted at college-bound high school juniors -- is having a substantial beneficial effect on senior course-taking behavior (Chronicle of Higher Education, June 25, 1986). But the prime intent of such programs is to ensure that all students are performing at or above a designated level on a set of common, predetermined performance dimensions.

A second path sees assessment primarily as a way of ensuring that institutions are making improvements along a number of performance dimensions deemed important by the state. These most commonly include effectiveness in providing general education, in delivering instruction in the major field, or in placing students in employment situations in their field of training. While common instruments are occasionally mandated by this approach, not every student need be tested, and a variety of approaches to testing are often used. This path is most clearly illustrated by Tennessee, and is the emerging dominant pattern for state-level assessment (for example, Maryland, New Jersey, and South Dakota). In Tennessee, for example, the Instructional Evaluation Schedule holds institutions accountable for their performance in five areas -- general education, achievement in the major field, student satisfaction, accreditation record of accreditable programs, and use of outcomes information to make actual improvements in programs and services. Only in the first is a common assessment approach mandated. On other dimensions, institutions may choose their own approaches. Indeed, the fifth dimension is not an outcome at all. Rather it is a desired institutional practice, and is evaluated in terms of the evidence provided that a satisfactory process is in place and is being utilized. In Colorado, no assessment methods are specified, and it is unclear as yet whether common statewide testing will occur. But enacted legislation does mandate the kinds of outcomes and practices that the legislature felt to be most important. Among them are such elements as self-confidence, persistence, empathy, leadership, and a range of other attitudinal and behavioral outcomes (Colorado General Assembly 1985).

The final approach sees assessment primarily as a device to promote institutional self-improvement. Under this approach, institutions are responsible for demonstrating to state authority that they have an adequate process for determining the degree to which they are attaining their own instructional goals. While answers to a common set of process questions may

be required, there is no requirement for institutions to use the same methods, or even to assess the same dimensions of performance. This approach, which somewhat resembles existing regional accreditation practice, is illustrated by Virginia. In Virginia, a strong initial predisposition is to preserve a healthy diversity among the state's current array of public institutions. Any approach that might tend to obscure well articulated and differentiated missions is to be avoided. As a result, Virginia's approach confines itself to requiring institutions to assess educational outcomes, but without defining a common set of assessment questions (Commonwealth of Virginia 1986). Such a process has much in common with existing mechanisms of statewide academic program review -- for example, in Illinois. As part of statewide Resource Allocation and Management Plan procedures, each public institution in Illinois must regularly assess the effectiveness and quality of its academic programs in terms of resources, enrollment, program need, and output. Institutions may design their own processes for undertaking such reviews but must report on the process to their own governing boards and to the state coordinating board. Indeed, one available state response to the call for assessment is to incorporate into established program review mechanisms explicit requirements to evaluate educational outcomes. Traditionally, as in the Illinois example, such processes have relied primarily on resource comparisons and peer review.

It is important to note that distinctions among these three approaches do not in themselves entail use of different methods of testing. All three, for example, may employ standardized instruments. "Rising Junior" testing programs such as those in Florida and Georgia involve the most visible use of standardized cognitive achievement tests. Tennessee's performance funding approach, however, while falling in the second category of state action, nevertheless mandates common use of the ACT-COMP instrument to assess each institution's contribution to undergraduate general education. Also in the second category, emerging programs in New Jersey and Maryland plan administration of common statewide tests to assess general education. In South Dakota and Tennessee, moreover, one suggested approach to assessing student achievement in the major field is administration of an available standardized instrument such as the Graduate Record Examination (GRE).

Furthermore, programs initially intended for one purpose may develop further once this purpose is achieved. This is amply illustrated by institutional experience with assessment programs. For Northeast Missouri State University, for example, a primary initial intent of assessment was to demonstrate the institution's competitive position with regard to student achievement (McClain and Krueger 1985). For this purpose, nationally normed cognitive tests were required, and the university's assessment program was built around ACT Assessment, ACT-COMP and GRE scores. With its credibility established, however, the university finds increasing utility in using the information gathered for internal improvement, and is experimenting with faculty-designed local examinations and other forms of assessment. Institutions like Alverno College have experienced the same process in reverse: originally designed as an integral part of the curriculum, assessment results have considerably enhanced institutional credibility with outside constituents (Mentkowski and Loacker 1985).

Finally, cutting across all three basic approaches is the question of incentives. Statewide programs can, and have, been founded on a range of incentives, both positive and negative. In Florida, for example, the legislature was careful to appropriate additional funds for institutions to implement statewide assessment -- so-called "CLASP/Gordon" funds. These were provided on a one-time basis to foot the considerable costs involved in establishing the program. Maintenance of the Florida program, however, tends to depend upon a different set of incentives. By making student test performance a condition of advancement, the program places a strong deterrent against both individuals and institutions ignoring its provisions.

Policy approaches that tie explicit funding decisions to the results of assessment are the most common forms of incentive. For example, under Tennessee's Instructional Evaluation Schedule, an institution can receive up to 5 percent in additional instructional funding if it performs well. The opposite approach is illustrated by Colorado; when the assessment provisions of House Bill 1187 come into force in 1990, institutions stand to lose up to 2 percent of operating budget if they do not appropriately comply. Regardless of the content of the program, the incentives behind it may have a profound impact on the degree to which it is successful.

This range of state approaches to assessment means that it is difficult to make a single estimate of probable consequences. Different types of programs will have different kinds of impacts, both positive and negative. The purpose of the following section is to discuss some of the probable impacts of each approach based on available evidence.

## IMPACTS OF DIFFERENT APPROACHES: SOME EVIDENCE

Evidence about the probable impact of state-mandated assessment is emerging from a number of sources. First, ten states (California, Colorado, Florida, Georgia, Maryland, New Jersey, South Dakota, Tennessee, Utah, and Virginia) have legislated or are otherwise in the process of establishing outcomes assessment efforts. While only two of these -- Florida and Tennessee -- are of long duration, lessons about the implementation of such programs are emerging constantly. Secondly, some two dozen institutions have implemented locally designed assessment programs on their own initiative over the past two years (see Appendix A). While few of these efforts have a long history, they share a number of common patterns of impact observed in several national projects (Ewell 1985b). A second "wave" of institutions is currently in the process of designing such programs, either in response to emerging state initiatives (or the expectation that these will develop) or to new regional accreditation criteria that stress collection and use of information on institutional effectiveness (SACS 1985). The Southern Association, for example, now requires all institutions to address the question of institutional effectiveness as part of the self-study process. Among the kinds of data it encourages institutions to collect are studies of student cognitive gain, attitudinal development, and job success after graduation. As evidenced by a recently completed survey, other regional accrediting organizations are currently moving to stress educational outcomes as a part of self-study (Thrash 1986). The author and other observers have been in direct

contact with the institutions in consulting roles and can credibly report on impact to date. Finally, higher education has in the past experienced a number of reform efforts that are structurally similar to state-mandated assessment programs. These include formal statewide processes for program approval and review, cost studies, and the professional accreditation process. Evidence of institutional response to such initiatives provides another resource for estimating the probable impact of proposed statewide assessment programs.

As previous discussion has emphasized, emerging state experience with assessment has tended to cluster around three major alternative approaches. These alternatives are: a state-run minimum skills assessment program in which individual students are tested primarily to provide quality assurance; a program designed to ensure that institutions are making changes in curriculum and instruction along designated dimensions of effectiveness consistent with statewide goals; and a program designed to encourage institutional self-examination consistent with each institution's own mission and instructional philosophy. Probable impacts of each of these approaches will be discussed in turn below.

State-Run Minimum Skills Testing

Because this is the most prescriptive of the three alternatives, opponents fear that its impacts will be the most severe. Foremost among claimed impacts is the fact that whatever mechanism is chosen to assess skills -- and this will undoubtedly be a standardized cognitive achievement test -- the tendency will be toward a narrowing of the curriculum and "teaching to the test." In practice, our observation indicate, it is necessary to temper this claim. Certainly there is considerable evidence that mandating performance on a single standardized instrument does indeed have some of these effects. At Miami-Dade Community College, for example, CLASP testing requirements have resulted in an increased focus on communication and computational skills in the curriculum (Losak and Kelly 1986). Moreover, our consulting work in Florida has surfaced reports that student time spent reading and writing has increased due to the state's minimum skills initiatives.

Such impacts have been strongly paralleled by experience at institutions that have relied heavily on standardized testing for evaluation -- most notably, Northeast Missouri State University (NMSU). After assessment evidence revealed that students were regressing in mathematical skills between the freshman and sophomore years, the university increased its emphasis on computation in general education courses, and eliminated a number of "soft-track" mathematics and science courses (McClain and Krueger 1985). Moreover, senior seminars at Northeast are partly developed as preparation for the Graduate Record Examinations, pre-professional tests, or professional certification examinations used as senior tests. While this has caused some concern about a "narrowed" curriculum, it also has produced impressive results on the tests themselves. For example, NMSU recently achieved the highest pass rate in the nation on the CPA examination largely through curricular tailoring of this sort.

Evidence from specialized accreditation of professional programs reinforces this perspective. One result of this process, criticized in several national reports (NIE 1984), is the tendency for specialized accreditation requirements to dominate course offerings in professional programs -- virtually crowding out any courses not directly related to occupational or professional training. If statewide assessment programs are designed only in terms of the mastery of a narrowly defined range of skills, similar results may be expected. On the other hand, stressing student attainment in general education -- particularly in basic skills and in such areas as critical thinking -- may provide institutions with needed incentives to fight the kinds of over specialization fostered by professional accreditation. At the very least, they may induce professional curricula to devote more attention to producing such outcomes in their own coursework.

In general, our observations of about forty institutions indicate, the major results of minimum skills testing have been that greater attention is paid to the sequencing of courses in a curriculum based upon necessary prerequisite skills, and greater attention is given to reinforcing basic skills such as reading, writing, and quantitative analysis across the curriculum. Whether such changes are uniformly positive, of course, depends upon one's point of view: from one perspective, such changes constitute "tightening" curriculum, while from other perspectives, such initiatives "narrow" it. In general education, for example, one tendency of assessment has been for institutions to reduce the number of courses acceptable to meet institutional general education distribution requirements. Certainly this reduces student choice, but the tendency has been to eliminate "soft-track" courses (particularly in the sciences and mathematics) in which college-level skills are not exercised.

The same can be said with respect to the impact of such testing on the conduct of instruction in the classroom. Certainly available evidence suggests that mandating a single, statewide cognitive achievement test results in focusing instructional effort on the content covered by this examination. The evidence is also clear that significant test score gains can result from this kind of instructional tailoring (Pace 1979, 1985). The question, of course, is whether or not some tests ought to be "taught to." If the purpose of assessment is indeed to ensure minimum skills without which a student cannot effectively progress or function in society, or if the assessment chosen is the officially recognized and rewarded method of entry into a particular profession (for example, medical boards or bar examinations), "teaching to the test" may well constitute a valid component of effective instruction. The important point is to avoid replacing traditional methods of student performance evaluation with minimum or advanced skills performance on standardized tests.

Minimum skills testing also raises the question of who can be legitimately held accountable for what. On the one hand, traditional academic practice holds individual students largely accountable for mastering course content. In this view, exposure -- or merely providing the curriculum as advertised -- effectively discharges the institution's responsibility. On the other hand, the fact that public institutions are explicitly chartered to help

meet identified public needs for manpower development and an educated citizenry suggests that institutions also need to be held responsible for the quality of their aggregate educational product. As a result, programs that exclusively test students, and that make students bear the cost of failure, have the potential of avoiding needed emphasis on institutional self-improvement.

These issues can interact in complex ways. For example, one regional public university experimenting with "rising junior" tests in mathematics as a condition for advancement discovered that student performance varied significantly by course section (Ewell 1985a). Such results made it obvious that consistent instruction in mathematics was not being delivered by the institution's faculty. From the standpoints of both equity and the law, it made little sense to hold students individually accountable for obvious inconsistencies in instruction over which they had no control. Rather, the need here was programmatic -- to use assessment results to help rectify these inconsistencies. Certainly if the primary purpose of assessment is ensuring that students have the requisite skills to engage in college-level work, entry testing of individual students is a legitimate exercise. If students also are tested on exit, however, some mechanism should be present to ensure that institutions bear some of the responsibility for addressing high levels of failure.

In practice, the evidence suggests, some such incentives are indeed present. First, because minimum skills performance results are comparative and are part of the public record, they undoubtedly affect institutions as they compete for students and for external funds. An enhanced competitive recruitment position, for example, has certainly been an outcome of assessment for institutions such as Northeast Missouri. Failure to maintain enrollment, of course, in states that have formula budgeting, will directly affect institutional funding. Student performance on assessment devices that may decrease or maintain enrollment therefore become of vital institutional interest. In one case, for example, a regional college discovered that it stood to lose almost a third of its upper-division enrollment if it enforced proposed standards for student achievement on a standardized mathematics test (Ewell 1985). It is now too early to tell if minimum skills testing will substantially affect enrollment patterns and therefore levels of funding across institutions. This potential linkage, however, constitutes a powerful incentive for institutions to monitor and improve their students' test-taking performance.

Moreover, results of Florida's CLASP testing and ACT-COMP in Tennessee are rapidly becoming the object of institutional comparison in the media. Such exposure certainly serves as a stimulus for action, though critics claim that the changes that result are constricting in the absence of information on the way institutions differ in terms of mission and primary clientele. In Florida, for example, the impact of CLASP testing on institutions with heavy minority enrollments has been much greater than on those enrolling fewer academically deprived students. Furthermore, direct impacts have been felt much more strongly among community colleges than among four-year institutions.

In sum, common minimum skills testing as a statewide initiative makes sense only if its intended consequences are clear. Foremost among these purposes is to ensure that individual students are indeed appropriately able to benefit from college-level work. Because of its prescriptive nature, the impact of such programs on institutions will be direct and considerable, and will indeed include such responses as a tailored curriculum and directed teaching. In the area of minimum social functioning skills, or the specialized skills needed to perform effectively in particular professions, such impacts are probably desirable. But states should carefully define and limit their efforts to such commonly defined or easily recognized outcomes. Common testing of all individuals is less effective if broader areas of knowledge and skill are to be assessed. Here the technology of assessment is much less reliable (Pace 1985), and greater freedom should be allowed for institutions to experiment with their own approaches. More importantly, the impacts of narrowed curriculum and tailored teaching that might be a virtue in producing basic and professional skills may be counterproductive in producing higher order skills. These issues are discussed in the following section.

## Assessment in Terms of Common Statewide Objectives

To date, about half of the established and emerging state assessment programs address common statewide objectives. In their most frequent form, for example, New Jersey, South Dakota, and Tennessee, programs of this sort require institutions to address a specified range of programmatic and institutional outcomes -- for example, student achievement in general education or the major field, student placement in their fields of training, or student satisfaction. While common instruments or methods are occasionally mandated by such programs -- particularly in general education -- in most cases institutions are encouraged to develop their own assessment procedures. It is expected that the approaches that they develop will speak to statewide priorities, while at the same time will reflect appropriate institutional differences.

Certainly experience to date suggests that there is likely to be little resistance to state-mandated choice of assessment dimensions -- particularly when such dimensions are fully discussed, and when adequate time is allowed for institutions to react and propose modifications. The evolution of the Tennessee Instructional Evaluation Schedule from a program that emphasized the particular outcomes to be attained, to one that currently stresses the kinds of assessment practices that institutions should undertake, is particularly instructive (Levy 1985). Several other states are currently engaged in such dialogue -- most notably New Jersey and Colorado. As long as institutions are allowed reasonable latitude in selecting appropriate methods of assessment within the state framework, little resistance occurs. Much more arises when a single test is mandated, even if its purpose is curriculum evaluation.

In addition to stressing institution-specific improvement, this approach emphasizes aggregate outcomes -- that is, the states that follow this approach do not mandate universal testing that has consequences for individual student promotion or graduation. Increases in individual student learning in the classroom are thus held to be indirect products of curricular reform. Certainly available evidence from institutional programs and from Tennessee

strongly supports the claim that assessment programs of this kind will have an impact on curriculum (Ewell 1984, Banta 1985b). Unlike universal testing, however, programs that stress self-assessment on mandated dimensions tend particularly to affect the clarity and integration of a curriculum, rather than its cognitive content (Ewell 1984).

One difficulty involved in examining the curricular impacts of any reform effort is knowing exactly what the term "curriculum" refers to. As Joan Stark (Stark and Lowther 1986) has pointed out, "curriculum" can refer to one of three quite different things -- what the catalog says, what the faculty actually teach, and what the students actually experience. Certainly one important impact of programmatic assessment as mandated by this approach is that it tends to bring these three elements together. For example, Alverno College describes its curriculum in outcomes terms: catalog descriptions of courses contain the competencies and performance skills that students are expected to master, and faculty assessment of these skills are built directly into classroom teaching (Mentkowski and Doherty 1984). At a recent planning retreat at Northeast Missouri, moreover, faculty curriculum planning was explicitly guided by concrete profiles of student knowledge and skill levels, and of attitudes toward learning, proposed by each departmental faculty. These planning profiles were anchored by actual past assessment results, and served as a visible template for designing curriculum structure.

Indeed, experience has shown that the need to design an assessment method in response to an external mandate can in itself be beneficial. At the University of Tennessee, Knoxville, for example, faculty in several departments revised the structure and content of introductory courses in their disciplines as a result of their efforts to design a satisfactory major field examination; explicitly deciding what a student needed to know at each point in a curricular sequence itself proved sufficient to stimulate needed change (Banta 1985a, Banta 1985b). At the same time, the CLASP effort in Florida has clearly stimulated several institutions to move beyond a minimum skills orientation. For example, both Miamia-Dade Community College and St. Petersburg Junior College are typical of Florida community college curriculum reforms; both institutions have largely reformulated curriculum and course descriptions in terms of expected outcomes.

Assessment results of this kind also can have a profound effect on course sequencing. For example, locally collected assessment information at North Carolina State University led to improved advising and placement of students in particular courses, and to uniform enforcement of existing prerequisite requirements; together, these efforts resulted in higher rates of student success -- particularly among minority students. Similarly, at Mt. Hood Community College in Portland, Oregon, use of information from locally administered placement surveys led to substantial changes in course sequencing in several occupational programs -- changes that led to higher success rates and greater graduate and employer satisfaction. In both these cases, assessment mechanisms were locally designed, but the impetus for assessment came from outside the institution. In North Carolina, this stimulus was provided by the university's need to respond to a federal desegregation Consent Decree, while follow-up data collection at Mt. Hood was in response to federal requirements for funding vocational education.

A parallel benefit is often visible at the level of the individual classroom. At Knoxville, many faculty report that the need to develop local field examinations as part of the assessment program has helped them to design better tests and performance assessments for use in their own classrooms (Banta 1985b). Furthermore, assessment can stimulate institutions to develop central resources to provide technical assistance to faculty who are seeking to evaluate course-specific performance. At the University of Tennessee, Knoxville, Alverno College, and Texas A&M University, such centers are increasingly used and appreciated by individual faculty members.

These curricular experiences are reinforced by evidence from states that have established formal processes for academic program review. Such processes resemble state-mandated assessment around a set of specified dimensions because they require all existing and proposed programs to respond to a common set of evaluative questions about projected enrollment, cost, need, and anticipated outcome. Despite their original intent, the major impact of such programs has not been to eliminate programs on the grounds of cost or unnecessary duplication. Rather, such programs have more often served as a stimulant for institutions to reexamine and sharpen curricular goals, to tighten course sequencing and curricular structure, and to eliminate as requirements courses not contributing to overall programmatic outcomes (Barak 1982, Conrad and Wilson 1985). It is important to note, however, that these experiences have been largely dependent upon program review being a highly participatory, faculty-centered process. Similar results can be expected (and indeed have been experienced) by assessment programs in which faculty participation is regular and meaningful. They are less likely if control for assessment is outside the institution, or if administration chooses to respond to the need for assessment without substantially involving the faculty.

As noted previously, a particularly important policy choice within this approach is whether common instruments or procedures are to be mandated. Several developing statewide programs, for example, prescribe a common instrument for assessing programmatic outcomes in general education. Using a common standardized instrument, of course, has the merit of allowing direct comparison of results among a variety of institutions. But such comparison can be risky if differences in institutional mission and clientele are not taken into account. Many states have experienced similar difficulties in their attempts to measure and compare program costs. Throughout the 1970s, many states calculated costs for delivering units of instruction by discipline and level, and compared the results across institutions. In doing so, they experienced difficulties because the clientele, mission, location, economy of scale, and other factors differed among institutional types, from major research universities to community colleges (Brinkman 1984). Over time, state agencies learned that appropriate cost comparisons are best made with sets of similar "peer" institutions, rather than across the state. Similar approaches are appropriate for comparing outcomes results. Indeed, in Tennessee, peer comparisons of ACT-COMP scores are an important facet of general education evaluation.

Furthermore, if common instruments are to be mandated, institutional experience with assessment suggests that they be of a particular character in

order to be effective in curricular improvement. First, the assessment instrument or technique chosen must be consistent with the curriculum as actually taught -- a requirement not often fulfilled by existing standardized test instruments (Turnbull 1985). Secondly, data from the instruments chosen must be fully available so that faculty (and students) can pinpoint areas of explicit strength and weakness (Banta 1985b). Again, this requirement is not often met by existing standardized tests that provide, at most, a few subscores in terms of which to summarize student performance. These requirements are of less importance if testing is intended to certify minimum competency. They are critical if test results are intended to inform curriculum or institutional improvement.

A second policy issue is whether to mandate more than a single source of assessment evidence. Most existing or emerging institutional programs involve at least three kinds of data collection -- standardized cognitive testing, locally designed cognitive testing, and student surveys. At Northeast Missouri, for example, assessment of general education involves the use of two standardized instruments (the ACT-COMP and ACT Assessment), and three student surveys. To comply with state assessment requirements, the University of Tennessee, Knoxville, like other institutions in the state, must integrate information from ACT-COMP, from standardized and locally developed field examinations, from a locally designed student survey, and from the professional accreditation process. Finally, Alverno College explicitly cites "triangulation" as a design requirement in its assessment process: all results are mutually substantiated by standardized testing, in-depth interviewing, locally designed task performance exercises, and external review of student performance by panels of practicing professionals. In general, experience has demonstrated that the more sources of evidence used in a given assessment, the less likely are the possibilities that curriculum will suffer or that faculty will teach to the test (Pace 1985). Encouraging multiple assessment practices also will help avoid what one college president has referred to as "the politics of the single figure" (Smith 1986).

In sum, statewide assessment approaches that stress the particular types of outcomes to be addressed can be of considerable benefit to both states and institutions. A clear specification of outcomes dimensions means that all institutions can be held accountable for answering the same kinds of questions. In areas such as general education and student satisfaction, investigation of this kind may only proceed adequately if the state asks the question; institutional experience has not in the past encouraged confidence in the willingness of faculty and administrators to ask such questions on their own (Astin 1985). But state authorities should simultaneously recognize that most of the benefits derived from such assessment programs depend upon the ability of the institution to adapt the measurement process to its own circumstances. This means allowing institutions considerable flexibility in carrying out the assessment mandate, while carefully holding them accountable for undertaking and maintaining a sound local process.

## Assessment as Institutional Self-Investigation

The third major approach to statewide assessment is to require institutions to undertake concrete, regular investigations of educational

outcomes in relation to their own instructional goals. Because this approach mandates neither what is to be assessed nor how it is to be assessed, some feel that it will be easily accepted by institutions, and will result in the same kinds of benefits as those experienced by institutions that have developed their own programs without such stimulation. In practice, however, experience suggests that it is usually not enough to simply require institutions to undertake "assessment" without describing more fully what particular aspects of performance they ought to cover. For example, much of the current debate over assessment in Colorado centers around the fact that House Bill 1187 mandates data collection on student outcomes, but in so unspecified a fashion that every institution's worst fears are stimulated. A similar process was typical when the Southern Association of Colleges and Schools first attempted to institute "institutional effectiveness" as a criterion for accreditation. Initial drafts of the standard aroused suspicion that much more was intended than was actually stated, and many institutions refused to cooperate as a result.

If indeed a state can induce individual institutions to undertake independent outcomes investigations, most of the impacts noted in the previous section also apply. Indeed, most of the evidence used in the previous section was drawn from assessment programs that were designed and carried out by institutions themselves. Such approaches, for example, are least likely to result in a narrowly specified curriculum and teaching to the test. Additionally, they are the most likely to yield substantial faculty involvement.

At the same time, such approaches may be least likely to induce meaningful institutional change. Under conditions of complete discretion to choose both assessment dimensions and the methods by which they will be measured, many institutions will claim that they already engage in meaningful assessment through the grading or curricular review process. More subtly, institutions may attempt to shift the burden of proof for educational outcomes onto the individual student. Some radical philosophical opponents of assessment, for example, appear to have fewer problems with universally applied minimum skills or achievement examinations than they have with approaches that hold the curriculum or instructional practice up for explicit critical examination. This is particularly the case if faculty have been solely responsible for setting minimum skill levels, or have designed the assessment instruments to be used. Indeed, low pass rates in particular courses have often been cited by faculty as evidence of programmatic rigor rather than being seen as an indicator of potential problems with what students are learning.

One way of forestalling this outcome is to require institutions to provide concrete examples of the ways they are actually using assessment information. One dimension of the Tennessee Instructional Evaluation Schedule, for example, is a review of the process by which institutions introduce assessment information into the academic decisionmaking structure. Even if no concrete assessment dimensions are prescribed at the state level, the reporting process should require institutions to demonstrate the mechanisms that they use to ensure that the information collected is making a difference.

A second problem with allowing institutions complete discretion regarding what to assess is that they may ignore important state-level priorities about undergraduate education. As long as institutional missions are differential and are effectively reviewed in terms of their contribution to the effectiveness of the state's higher education system as a whole, this may not be a problem. In Virginia, for example, the concept of system effectiveness being maintained by the Virginia Council encourages each institution to see its own role in relation to systemwide objectives. Without such a central focus for institutional missions, however, many regional public institutions will conceive of their teaching mission as similar to that of the flagship institution to whose status they aspire (ECS 1986). This may well induce such institutions to shift assessment priorities to emphasize such outcomes as graduate school placement and GRE examination performance -- outcomes that may not be appropriate given their clienteles and presumed teaching objectives.

In sum, most of the impacts of assessment on curriculum and teaching practice mentioned in the previous section also hold for this approach. State policymakers should be fully aware, however, that meaningful change under such an approach will depend upon the degree to which the institutions in a state system are clear about their own instructional missions, and upon the existence of a mechanism to ensure that their individual missions are consistent with statewide goals.

## SOME IMPLICATIONS FOR ACTION

Based on the above evidence, a number of recommendations for state action are appropriate. Because the range of state approaches is considerable, some of these recommendations apply only to particular types of state action, while others can be applied to any state-mandated initiative. They are presented under appropriate headings below.

If states contemplate universal minimum skills testing, they should:

- o Make sure that the purposes of the program are clear. Institutions should receive unambiguous signals that the intent of the program is narrow, and concerns readiness for different stages of college-level work. As a result, the range of outcomes tested should be explicitly limited to basic academic or social functioning skills. It should carefully avoid such areas as critical thinking, leadership, and other less specifiable outcomes.

- o Make sure that proposed testing does not replace existing mechanisms for evaluating and promoting students. While scores on such tests can legitimately serve a "gatekeeping" role for entry into different stages of college-level work, care should be exercised that degrees or promotion are not awarded solely on the basis of such performance. Nor should institutions be allowed to subtly drift into such a situation.

- o Find ways to make institutions as well as individuals accountable for results. Such programs should actively promote institutional responses to detected problems, either through direct incentive or through an explicit, periodic review or reporting process.

If states contemplate an approach based on institutional self-assessment in terms of common statewide objectives, they should:

- o Keep the number of assessment dimensions manageable and clearly link these dimensions to areas of legitimate state concern. General education and major field outcomes, for example, may be legitimately claimed as state concerns because of statewide goals for manpower development and an educated citizenry. Other areas, for example, desired value changes, may be less justifiable.

- o Reward the quality of institutional participation in the process, as well as the outcomes produced. Especially in the early stages of such programs, institutions will be reluctant to cooperate. Incentives should reward participation in the process, in addition to actual institutional performance.

- o Mandate common, standardized approaches only when there are clear benefits of comparison. A major virtue of such testing is that results can be compared both across institutions and with national norms. But such comparisons are universally valid only on certain dimensions. For general education, for example, a case can be made for common testing. Even here, however, comparison should be approached cautiously and should emphasize comparison with identified peer institutions rather than being statewide.

- o Allow institutions to focus their assessment efforts around legitimate differences in curriculum. Because curriculum structure and delivery will be institution-specific, the best approaches to improvement involve institutions designing their own processes. While flexibility is important, however, the process should clearly encourage better statements of curricular goals and better articulated program structures.

If states contemplate an approach based on institutional self-assessment around institution-specific goals, they should:

- o Make sure that the institutional missions that drive self-assessment are themselves clearly stated and periodically reviewed in a statewide context. Self-assessment is consistent with statewide goals only if "quality" is seen the same way from the state and the institutional perspective. Left completely to themselves, institutions may choose assessment dimensions and methods not suited to their originally assigned mission, or inappropriate to their current clientele.

- o Require institutions to report the ways they are using assessment results and the changes they have made in response to them. The program should clearly signal that the responsibility is for demonstrated improvement rather than for providing information. As a result, such programs might legitimately include explicit follow-up on proposed action in future years.

- o Ensure that the institution itself takes responsibility for outcomes, rather than placing the responsibility solely on students. It should not be possible to meet requirements for self-assessment merely by changing the ways in which individual students are evaluated. Results also should critically examine curricular and institutional practice.

Finally, in developing any statewide approach to assessment, the following recommendations apply:

- o Require institutions to report not only information, but how they will use it to make improvements. Institutions should be able to identify both the concrete mechanisms for using information and, over time, should be able to provide actual examples of improvement.

- o Stress positive incentives in implementing assessment. The best approaches provide direct, financial incentives -- as in Tennessee. Results also might be used to allocate "special fund" dollars for innovation or instructional improvement. Approaches that emphasize loss of funding, programs, or flexibility should be avoided.

- o Encourage the use of multiple indicators of performance. The best programs involve data collection on several dimensions, and according to several different methodologies. Institutions should be encouraged to avoid, and state mechanisms also should avoid, "the politics of the single figure."

- o Fully integrate requirements for assessment with other state mandates for improvement. Mandating assessment in isolation of other policy initiatives -- for example, "centers of excellence," new faculty reward structures, and challenge grants -- makes little sense. Instead, assessment policy should be used to help evaluate and shape efforts at reform in other areas.

- o Recognize that flexible approaches are better than rigid ones, but that specific approaches are better than ambiguous ones. Any statewide approach to assessment should recognize important differences among institutions with regard to instructional mission and primary clientele. But these differences should not lead to such broadly worded requirements that institutions are left fearful about what will and will not be covered.

- o Allow sufficient time for effective programs to develop. One important aspect of successful practice at both the institutional and state levels has been the opportunity for parties at interest to fully negotiate their differences, and to pilot-test proposed programs. Rushing to implementation may well damage institutional cooperation -- the primary determinant of whether state-mandated assessment will be successful.

Permeating all these recommendations is a single theme: statewide approaches should as far as possible be designed to challenge excellence

rather than to mandate adequacy. Experience has shown that higher education institutions are at their best when pursuing the best. They are at their worst when they are asked to respond to narrowly specified definitions of performance supplied from outside the academic arena. In challenging institutions to better articulate and maintain their own standards, state authority should be fully aware of both the opportunities and limits that these conditions imply.

## REFERENCES


American Association of Colleges. "Integrity in the College Curriculum: A Report to the Academic Community." Washington, D.C.: AAC, 1985.

American Association of State Colleges and Universities and Northeast Missouri State University. "In Pursuit of Degrees with Integrity: A Value-Added Approach to Undergraduate Assessment." Washington, D.C., 1984.

Astin, Alexander W. "Achieving Educational Excellence." San Francisco, Calif.: Jossey-Bass, 1985.

Banta, Trudy W. "The Use of Outcomes Information at the University of Tennessee, Knoxville." In P.T. Ewell (ed.), "Assessing Educational Outcomes." New Directions for Institutional Research, Vol. 47. San Francisco, Calif.: Jossey-Bass, 1985a, pp. 19-32.

Banta, Trudy W. (ed.). "Performance Funding in Higher Education: A Critical Analysis of Tennessee's Experience." Boulder, Colo.: National Center for Higher Education Management Systems (NCHEMS), 1985.

Barak, Robert J. "Program Review in Higher Education: Within and Without." Boulder, Colo.: NCHEMS, 1982.

Bennett, William J. "To Reclaim a Legacy: A Report on the Humanities in Higher Education." Washington, D.C.: U.S. Government Printing Office, 1984.

Bogue, E. Grady, and Brown, Wayne. "Performance Incentives for State Colleges." Harvard Business Review, Vol. 59, November/December 1982, pp. 123-128.

Brinkman, Paul T. "Formula Budgeting: The Fourth Decade." In "Responding to New Realities in Funding." New Directions for Institutional Research, Vol. 43. Larry Leslie (ed.). San Francisco, Calif.: Jossey-Bass, 1984, pp. 21-44.

Brinkman, Paul T., and Allen, Richard. "Concepts of Cost and Cost Analysis for Higher Education." AIR Professional File No. 23, Association for Institutional Research, Spring 1986.

Chaffee, Ellen Earle. "After Decline What? Survival Strategies at Eight Private Colleges." Boulder, Colo.: NCHEMS, 1984.

Colorado General Assembly. "House Bill No. 1187 Concerning the Reorganization of Higher Education." Denver, Colo., 1985.

Commonwealth of Virginia. "The Measurement of Student Achievement and the Assurance of Quality in Virginia Higher Education." Report of the Study conducted by the Council of Higher Education for Virginia to the Governor and the General Assembly of Virginia. Senate Document No. 14. Richmond, Va: Commonwealth of Virginia, 1986.

Conrad, Clifton F., and Wilson, Richard F. "Academic Program Review: Institutional Approaches, Expectations, and Controversies." Report No. 5. Washington, D.C.: ASHE/ERIC Higher Education Reports, 1985.

Education Commission of the States. "Time for a Different View: Transforming the State Role in Undergraduate Education." Report of the Working Party on Effective Action to Improve Undergraduate Education. Denver, Colo.: ECS, August 1986.

Ewell, Peter T. "Effectiveness Information and Institutional Change: An Exploratory Analysis." Boulder, Colo.: NCHEMS, November 1985a.

Ewell, Peter T. "Using Student Outcomes Information in Program Planning and Decisionmaking." Final report to the W.G. Kellogg Foundation. Boulder, Colo.: NCHEMS, November 1985.

Ewell, Peter T. "The Self-Regarding Institution: Information for Excellence." Boulder, Colo.: NCHEMS, 1984.

Leslie, David W. "Academic Freedom for Universities." Review of Higher Education, Vol. 9, No. 2, 1986, pp. 135-159.

Levy, Robert A. "Development of Performance Funding Criteria by the Tennessee Higher Education Commission: A Chronology and Evaluation." In Trudy W. Banta (ed.), "Performance Funding in Higher Education: A Critical Analysis of Tennessee's Experience." Boulder, Colo.: NCHEMS, 1985.

Losak, John, and Kelly, Terrence. "Testimony Before the Task Force on College Quality." Washington, D.C., 1986.

McClain, Charles J., and Krueger, Darrell W. "Using Outcomes Assessment: A Case Study in Institutional Change." In P.T. Ewell (ed.), "Assessing Educational Outcomes." New Directions for Institutional Research, Vol. 47. San Francisco, Calif.: Jossey-Bass, 1985, pp. 33-46.

Mentkowski, Marcia, and Doherty, Austin. "Careering After College: Establishing the Validity of the Abilities Learned in College for Later Careering and Professional Performance." Milwaukee, Wis.: Alverno Publications, 1984.

Mentkowski, Marcia, and Loacker, Georgine. "Assessing and Validating the Outcomes of College." In P.T. Ewell (ed.), "Assessing Educational Outcomes." New Directions for Institutional Research, No. 47. San Francisco, Calif.: Jossey-Bass, 1985, pp. 47-64.

National Education Association. Testimony Before the Task Force on College Quality. Kansas City, Mo., May 1986.

NIE Study Group on the Conditions of Excellence in American Higher Education. "Involvement in Learning: Realizing the Potential of American Higher Education." Washington, D.C.: U.S. Government Printing Office, 1984.

Pace, C. Robert. "Perspectives and Problems in Student Outcomes Research." In Peter T. Ewell (ed.), "Assessing Educational Outcomes." New Directions for Institutional Research, Vol. 47. San Francisco, Calif.: Jossey-Bass, 1985.

Pace, C. Robert. "Measuring the Outcomes of College." San Francisco, Calif.: Jossey-Bass, 1979.

Smith, Hoke L. Testimony to the National Governors' Association Task Force on College Quality on Behalf of the American Association of State Colleges and Universities, Washington, D.C., February 1986.

Southern Association of Colleges and Schools Commission on Colleges. Criteria for Accreditation. Atlanta, Ga.: SACS, 1985.

Stark, Joan S., and Lowther, Malcolm. "Curricular Integration: Student Educational Goals and Educational Outcomes." Ann Arbor, Mich.: Center for the Study of Higher Education, 1986.

Thrash, Patricia R. "Outcomes Measurement and Public Accountability." Paper presented at COPA Professional Development Session, Council on Postsecondary Accreditation, Washington, D.C., April 1986.

Turnbull, William W. "Are They Learning Anything in College?" In Change, November/December 1985, pp. 22-26.

# APPENDIX A

Institutions with Outcomes Assessment Programs
that NCHEMS has Worked With

| | | | | |
|---|---|---|---|---|
| Alverno College | Milwaukee | WI | private | 4 yr. |
| Northeast Missouri State Univ. | Kirksville | MO | public | 4 yr. |
| Univ. of Tennessee, Knoxville | Knoxville | TN | public | research |
| State Univ. of N.Y., Albany | Albany | NY | public | research |
| State Univ. of N.Y., Plattsburgh | Plattsburgh | NY | public | 4 yr. |
| James Madison University | Harrisonburg | VA | public | 4 yr. |
| St. Petersburg Jr. College | Tampa/St. Pete | FL | public | 2 yr. |
| Miami/Dade Community College | Miami | FL | public | 2 yr. |
| Kings College | Wilkes-Barre | PA | private | 4 yr. |
| Olivet Nazarene College | Kankakee | IL | private | 4 yr. |
| Mount Hood Community College | Gresham | OR | public | 2 yr. |
| Xavier University | New Orleans | LA | private | 4 yr. |
| Tenn. Technological University | Cookeville | TN | public | 4 yr. |
| Ohio University | Athens | OH | public | 4 yr. |
| Comm. College of Philadelphia | Philadelphia | PA | public | 2 yr. |
| Kapiolani Community College | Honolulu | HI | public | 2 yr. |
| Spoon River College | Canton | IL | public | 2 yr. |
| Southwest Texas State Univ. | San Marcos | TX | public | 4 yr. |
| Colorado State University | Ft. Collins | CO | public | research |
| Univ. of Colorado, Boulder | Boulder | CO | public | research |
| North Carolina State Univ. | Raleigh | NC | public | research |
| Towson State University | Towson | MD | public | 4 yr. |
| Kean College | Union | NJ | public | 4 yr. |
| Montana State University | Bowman | MT | public | 4 yr. |
| Rhode Island College | Providence | RI | public | 4 yr. |

# THE COSTS OF ASSESSMENT

Prepared by Peter T. Ewell, Senior Associate
and
Dennis P. Jones
National Center for Higher Education
Management Systems

The report of the Study Group on the Conditions of Excellence in American Higher Education, "Involvement in Learning: Realizing the Potential of American Higher Education" (1984), identified assessment and feedback as one of three conditions for achieving excellence in undergraduate education. The Study Group argued that "institutions should be accountable not only for stating their expectations and standards but for assessing the degree to which those ends have been met" (p. 21). The underlying theme is that acquiring and using information about performance is a necessary ingredient in any attempt to foster learning and self-improvement. We all recognize the legitimacy of this argument when applied to students. Most accept the notion that such evaluations need to be formative as well as summative. The assessment process may be implemented badly at times, but there is substantial agreement that the evaluation of student learning and development ought to guide the teaching and learning process.

Both the Study Guide and Peter T. Ewell (1984) go further. They argue that what holds true for assessing students also holds true in a broader context. Specifically, they maintain that the road to improvement of courses, programs, and indeed the institution itself, involves regularly collecting information on institutional and program effectiveness, and using such information as the basis for improvement. Intellectually and conceptually, the argument has the ring of reason. We can readily accept the notion that information in knowledge, and that we ought collectively to be more knowledgeable about our institutions and programs they house. Ergo, assessment information about institutions and programs, as well as individual students, is desirable.

On a more practical level, however, the recommendations of "Involvement in Learning" with regard to assessment and feedback are often greeted with skepticism. Indeed, the level of skepticism itself is revealing; it stems primarily from unfamiliarity rather than from unfortunate experience. The skepticism that we have observed surfaces in the form of two concrete questions. First: "Can assessment actually be accomplished; is it feasible?" As a technical question, this is being answered in the affirmative, supported by a growing body of institutional experience with wide-ranging assessment programs. But the second question is equally pragmatic: "How much does it cost?" The underlying tone of the question reflects a conviction that the costs are high.

We address the latter question in this paper. In the following section, we present a simple conceptual schema to delimit the dimensions of the question. In the balance of the paper, we present estimates of the costs of assessment for different types of institutions.

## AN ANALYTIC FRAMEWORK

To properly address the costs of assessment, we must pose and answer two distinct questions. The first question is "the costs of assessing what?" -- a

---

Reprinted with permission from the Office of Educational Research Improvement. "Assessment in American Higher Education: Issues and Contexts." Cliff Adelman (ed.) with forward by Secretary William J. Bennett. Washington, D.C.: U.S. Department of Education, 1986, pp. 33-46.

question of unit of analysis. The second question is "what costs?" -- a question of what to count. These two questions are treated separately below.

## Unit Of Analysis

The unit of analysis with which we are most traditionally comfortable in assessment is the individual student. In the normal course of events, the individual student experiences a wide variety of assessments in the process of being admitted to, and making progress through, an institution. Students take ACT and SAT tests as part of the application process. Incoming freshmen commonly take a battery of institutional tests for placement purposes immediately upon arriving at the campus. Most pervasive of all assessment activities are the many tests that students take in each and every course in which they are enrolled. By such means, we collect mounds of assessment data on students. Our facility for turning this data into information, however, remains limited. But we do at least gain enough information from these activities to convince ourselves that individuals do or do not deserve to be certified as academically worthy and eligible to receive a degree, diploma, or certificate.

Beyond the individual student, the units of analysis with which we are primarily concerned are the program or curriculum, and the institution as a whole. With regard to individual programs or curricula, assessment questions abound. A central question is: "Are the students who have completed the program emerging with the intended level of knowledge and skills, and are they proceeding to fill intended roles in desirable ways?" Corollary questions include the attractiveness of the program to particular groups of students and student satisfaction with the educational experience provided by the program. Each of the questions can be illuminated by periodic assessment of the outcomes of the program. While many of the basic data needed to address these questions are the same as those needed to assess individual student development, the ways in which we analyze these data will be different. For program evaluation, the primary need is to look at the collective performance of a particular body of students (or a representative sample thereof). This means not only examining mean or median performance, but also investigating and accounting for the nature of variations around these central tendencies.

Finally, comprehensive assessment requires information about the performance of the institution as a whole. It is at this level that questions of feasibility become most widespread and acute. As a consequence, it is at this level that assessment is least frequently conducted. Given a wide array of outcomes attributable to almost any college or university -- and given that the typical institution tends to claim credit for contributions to growth along all of these dimensions -- there is an understandable inclination, in Kenneth Mortimer's words, to "measure everything that moves." With this perspective, it is easy to see how questions of cost emerge as a real issue.

To estimate the costs of assessment, we first deal with the appropriate scope of assessment. With this requirement in mind, we want to emphasize that the essence of institutional assessment is to "measure your mission." Adherence to this simple principle can help insure that institutional

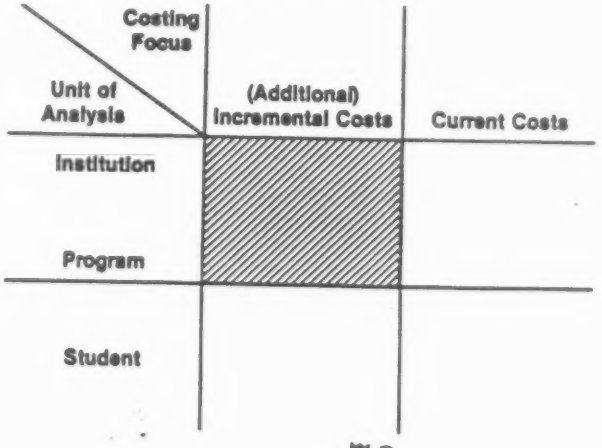
assessment is a carefully focused activity. Posing the question in this way also requires that assessment be tailored carefully to reflect the distinctive aspects of each institution. If the institution in question is primarily oriented toward professional and occupational training, appropriate assessment should be focused upon the documented success of graduates in the professions and occupations for which they were trained. For liberal arts colleges, in contrast, primary assessment strategies tend to examine student development along dimensions of general knowledge and general skills. In major research universities, assessment may be concentrated on student success in the major. There are, of course, variations on all of these themes, including consideration of student satisfaction with the experience, as well as educational "value-added."

## The Costs Considered

There are innumerable concepts of and ways to calculate, costs. Among them are direct costs, indirect costs, full costs, average costs, marginal costs, and opportunity costs. The appropriateness of each of these approaches to costing is determined by the use of the resulting information. Consequently, the real question for us is not simply "What is the cost of assessment?" Because the issue is usually raised in a managerial or resource allocation context, our question becomes, "How much more money do we have to spend to put in place an assessment program that is appropriate to our needs?" Using this notion as a guide, we have passed over attempts to estimate the cost of student assessments already undertaken as a regular part of the student's coursework. While it might be possible to calculate the actual proportion of faculty instructional effort already attributable to in and out of class assessment activities, this would yield information without a purpose. It is more important to attempt to determine the level of regular investment the institution must make in addition to these ongoing activities. This in an incremental or marginal cost. Certainly we recognize that the dollars currently being spent for assessment can often be spent more effectively, and that assessment programs can often be improved at no added cost. Such reallocation issues, however, are not within the domain of this paper.

Putting these two dimensions together results in a matrix that displays major cost considerations as follows:

DIMENSIONS OF TOPIC



The portion of assessment costs we discuss in the balance of this paper is indicated by cross hatched areas in this diagram. In the following sections, we provide estimates of typical incremental costs for establishing and maintaining institutional and program level assessment programs. At best, these estimates are exceedingly rough. In spite of their limitations, however, they do provide reasonable ballpark figures regarding the level of costs that might be expected by an institution embarking on a comprehensive assessment program.

## ESTIMATING COSTS OF AN INSTITUTIONAL ASSESSMENT PROGRAM

Estimating the actual new costs of establishing an assessment program at a given college or university can be a complex undertaking. As most institutions already collect some data on student achievement and program effectiveness, creating a comprehensive assessment program may thus involve coordinating a number of activities for which the institution has already paid. An additional difficulty is the level of analysis at which assessment takes place. While data gathering on institutional effectiveness remains relatively rare, all institutions collect some data on individual student performance. As noted above, the kinds of data routinely collected on individual students at most campuses may or may not be consistent with good depagogy. But in general, changing assessment methods and policies at this level will not entail significant additional costs.

Because of these difficulties, several caveats are necessary before we present some actual cost estimates. First, we will base our estimates primarily on direct costs -- those costs incurred by fielding new test and survey instruments, and by making use of the results. While a variety of indirect or overhead costs might be considered (for example, professional time spent drawing the implications of assessment results, faculty and administrative time spent reviewing programs in light of assessment data, and the like), these will vary so greatly that concrete estimates would be problematic.

Our second assumption is that an institution will adopt an explicit program for assessing instructional effectiveness. This means that various related instructional evaluation efforts will be centrally coordinated and supported by a staffed, visible office. Establishing such an assessment program, it is important to note, may involve considerable reallocation of existing, funded functions. For example, many institutions already fund a testing center, an institutional research office, or an academic planning office. Functions of each of these existing offices are commonly included in a comprehensive assessment program. Furthermore, many individual data gathering efforts included in a comprehensive assessment programs may already be in place in one or more of these locations. Many institutions, for example, regularly administer student surveys, such as the Cooperative Institutional Research Program (CIRP), or conduct surveys of students upon graduation or withdrawal. Many institutions regularly assess student abilities for placement purposes on entrance. Finally, many institutions regularly administer professional or pre-professional certification tests that assess particular skills gained in the course of instruction.

Our final assumption is that assessment will rest primarily on traditional "paper-and-pencil" testing and survey methods. Certainly there are many alternatives to cognitive tests and forced-choice survey instruments, and they should be carefully considered in building an assessment program. Use of external assessors drawn from the local business and professional community, as practiced by Alverno College and others, constitutes one such alternative (Mentkowski and Doherty 1983). Traditional jury or panel ratings of performances in such fine arts disciplines as music, drama, and dance provide another. In such cases, assessment costs can be estimated in terms of the professional time committed by external evaluators. The issue of costs is far more difficult when faculty themselves play these roles -- either in addition to, or instead of, traditional grading practices. In such cases, the "costs of assessment" can easily be viewed as part of an individual faculty member's existing assignment. Because of these difficulties, we consider only the direct costs of more traditional assessment methods in the discussion that follows.

## Cost Elements for Assessment Programs

In constructing assessment programs, most institutions incur costs in four basic areas. First, assessment instruments (tests and surveys) must be constructed locally or purchased from an outside vendor. Second, these instruments must be administered to students. Third, the resulting data must be analyzed and disseminated. Finally, the assessment effort itself must be coordinated. Each of these costs is driven by different parameters, and by the kinds of choices that institutions may make within each cost element.

### Instrument Costs

Various cognitive tests and students surveys form the basis for any assessment data gathering effort. Before they can be fielded, tests and instruments must first be developed or obtained. If they are developed locally, costs are incurred by faculty and measurement specialists in constructing the test or survey. After initial development, such instruments can be produced on a regular basis -- generally at lower cost than comparable commercial instruments.

The alternative to constructing instruments locally is to make use of commercially available tests and surveys. Examples include the Graduate Record Examination (GRE) Field Examinations -- often used as senior assessments of knowledge in the major field, various professional and pre-professional certification and placement tests (for instance, the National Teacher Examination) used for the same purpose, and course content examinations such as the College Level Examination Program (CLEP) -- designed for awarding credit-by-examination, but increasingly used to assess mastery of lower-division basic skills. Examinations such as these are obtained through purchase -- generally on a per-instrument basis.

Because of the difficulties involved, relatively few institutions choose to design their own cognitive tests. Generally, local achievement tests are developed as senior assessments in fields not currently covered by existing commercial instruments. More rarely, colleges have developed their own

general education assessment instruments. Local examinations have also been developed because faculty feel that existing commercial instruments do not adequately cover the field as taught in their own curricula. Developing good subject area examinations can be a time-consuming exercise and additional resources are required for pilot testing the instrument and for subjecting individual test items to careful review by testing/measurement specialist.

Institutions that have constructed such field examinations have usually treated their development as a departmental activity. This practice tends to bury many test-making expenses in ongoing departmental administrative costs associated with curriculum development and review. If the full cost of such activity were calculated, it would undoubtedly be quite high. In practice, however, budgeted test development costs tend to be treated as a short-term overload assignment for particular departments -- covering part, but not all, of the resources required. For example, one major research university is currently undertaking development of twenty such departmental examinations at a budgeted cost of $2,000 each. This university judges the cost as an appropriate increment only because departments are expected to reallocate additional existing resources to test-making that are already "budgeted" for curriculum review and improvement activities.

Development of local surveys -- either of currently enrolled students or of former students (graduates and dropouts) -- is much more common than development of local cognitive assessment instruments. In general, good survey instruments can be designed for less than the costs associated with cognitive tests. Some economies result from the fact that many common models are available. References such as McKenna (1983), Pace (1975), and California Community Colleges (1984) provide excellent and accessible lists of items commonly included on student surveys.

Commercial tests and surveys are generally purchased on a per unit basis. For cognitive tests and examinations, the unit price includes scoring as well as the price of the instrument. Individual prices vary considerably from a low of $7 per exam for instruments such as the ACT Assessment Entrance Examination, through $29 per exam for the GRE, to a high of $43 per exam for such instruments as the National Teacher Examination administered by ETS. In some, but not all, cases multiple purchase discounts are available for institutions.

Commercial student surveys are generally available for individual purchase, with or without associated processing and analysis services. Prices for individual instruments range from a low of 15 cents per survey to approximately $1 per survey. When analysis services are used, total costs average $3 to $5 for each completed questionnaire. In addition, institutions can purchase a tape of responses for $40 to $150 and can obtain comparative reports consisting of responses from other institutions that have used the instrument.

Administrative Costs

Once in hand, tests and surveys must be administered to students. In some institutions, existing testing centers established for placement or diagnostic

testing may bear some of this burden. In most cases, however, the number of instruments to be administered simultaneously will require resources beyond those available to the typical institutional testing center. Cognitive test administration is generally a straightforward, in-class exercise, but even so, considerable administrative costs may be incurred. For cognitive tests, proctors may need to be employed for multiple test locations.

For some types of tests (for example, the ACT College Outcomes Measures Project) special video and audio equipment must be available and operated. If special testing sessions are scheduled, students must be notified where they should appear, and follow-up procedures put into place to insure that they appear. Finally, costs will be incurred in recording results, and if desired, in sending test results directly to students.

Some of the same procedures will be typical of in-class survey administration. Generally, however, in-class student surveys will not require supervision, and survey questionnaires will take less time to complete than examinations (an average of 10 to 20 minutes as compared to the typical three-hour length of most examinations). Moreover, many student surveys can be administered in existing settings, for example, at student registration or during orientation programs. Because of the ready availability of such mechanisms, entering student questionnaires are the kinds of survey instruments most easily and directly administered to students.

For program graduates or withdrawing students, or for currently enrolled students who may be difficult to reach in an available "captive" setting, survey administration by mail will be typical. Mailed survey costs vary with the number of respondents to be reached, the number of mailings undertaken to maximize response, and the estimated response rate. In order to obtain acceptable response rates, most institutions use more than one mailing, and often supplement results with telephone follow-ups of non-respondents. Most sources recommend the use of first-class postage on both mailout and return envelopes (Dillman 1982). Costs for recording and tabulating responses should also be included in any analysis. Based upon such parameters, typical costs for conducting mailed surveys will average $1.50 to $3 per completed instrument.

Analysis Costs

As noted above, commercial cognitive tests include analysis and processing expenses with the cost of the instrument. Scoring and analysis services are also available for most commercial surveys. These services include costs for data entry, computer analysis, and production of a simple frequency or cross-tabulation report. In many cases, however, available data will need to be further analyzed for policy purposes. In the case of test data, individual student performance results may be correlated with student characteristics, with course-taking patterns, or with other elements of the institutional experience. This task entails creating data sets which make use of a variety of data elements beyond simple test performance. The same is true of student survey data. Tapes of questionnaire responses -- generally available from the providers of the instruments -- can also be analyzed locally using an available statistic package. All such exercises entail both personnel and data processing costs.

In the case of locally developed tests and surveys, analysis designs will have to be created from scratch. Like instrument design, this is a one-time cost, but it can be considerable. A set of analysis routines must typically be written using a standard statistical package (for example, SPSS or SAS) or using a common programming language.

Similarly, response coding schemes must be devised and, if applicable, matching scoring procedures using mark-sense equipment established. In the initial stages, considerable care must be taken to develop error-checking procedures and methods for handing missing, incomplete, or contradictory information. Once such procedures are put in place, however, ongoing costs for data analysis will be minimal, involving personnel and computer time.

Coordination Costs

Establishing a comprehensive program of institutional assessment may require investments beyond the direct costs associated with procuring, administering, and analyzing a variety of data gathering instruments. Such comprehensive programs are centrally administered and involve coordinating many kinds of data collection and analysis activities. Indeed, the most effective of such programs are located in distinct, specially created offices -- for example, Alverno College's Office of Research and Evaluation, and the University of Tennessee at Knoxville's Learning Research Center.

Costs associated with establishing an office of this kind include those for new professional and support staff, office space to house these personnel, and ongoing operating expenses. In estimating such costs, it is important to attempt to isolate the new functions that such offices will fulfill from those associated with the existing, previously funded activities that such offices will now assume. For example, overseeing an annual alumni survey effort and writing data reports on this activity may already be part of an institutional research responsibility and may be built into the cost structure of a new assessment office. Similarly, existing diagnostic testing and measurement activities may be folded into such an office's established responsibilities. Generally, however, such functions as administering comprehensive examinations in general education, or working with faculty to develop local survey and test instruments, are not covered in the institution's current cost structure.

In many cases, existing personnel are reassigned to provide staffing for an assessment center. Faculty with appropriate research backgrounds in the social and behavioral sciences, or Testing/Institutional Research professionals, may be taken from their current assignments and given responsibility for coordinating institution-wide assessment activities, for designing instruments, or for analyzing and interpreting test or survey results. In such cases, estimating costs may be difficult because the relevant question is the cost of replacing the reallocated staff member in his or her original function. Often this can be done with part-time instuctors or research assistants at a cost that is probably far less than that of full replacement. In other cases, the reassigned person may be currently underutilized, and may consequently not need full replacement. Alternatively, however, the reallocated position may be in a high-demand area, and a premium must be paid for its replacement.

As a result of the extreme variation in current practice, any estimates of coordination costs will be approximate. In each of the cases discussed below, we attempt to disaggregate these costs so that only the new costs associated with establishing an assessment program are counted. When these costs involve reassignment of existing personnel, the full cost of replacement provides the basis for the estimate.

## CONSTRUCTING TAILORED INSTITUTIONAL COST ESTIMATES

Because institutions vary widely in size, programs, and clientele, appropriate assessment programs will vary as well. A small, private, residential, liberal arts college will probably emphasize general education in its instructional mission to a degree not typical of a community college or a large research university. Consequently, it will appropriately concentrate its data gathering and analytical resources on assessing liberal learning outcomes. In contrast, a community college most likely will concentrate the design of its assessment program on job success and transfer to senior institutions. The clientele of the small liberal arts college will be much more conducive to administering tests and surveys in classroom settings than will be the case for the more dispersed community college population. As a result, methods for administering tests and surveys will vary considerably among types of institutions.

For illustrative purposes, we have constructed typical assessment programs for four types of institutions. They include: (1) a private liberal arts college with a traditional, residential student population of approximately 1,000 students; (2) a major public research university with a total student population of approximately 25,000 students (including 18,000 undergraduates); (3) a regional, comprehensive, public university with approximately 5,000 residential and commuter students; and (4) a mid-sized community college with an enrollment of approximately 15,000 students in occupational, transfer, and community service programs.

For each institution, we produced cost estimates as follows. First, based upon presumed instructional mission, we made a choice about which assessment dimensions should be emphasized. Second, we selected a typical array of instruments for each case, and estimated the direct costs for instrument procurement using published cost data for commercially available instruments, and common institutional experience for locally constructed instruments. Third, we chose a set of administration and analysis methods based upon expected student characteristics. Finally, we estimated coordination costs on the basis of the experience of existing data gathering and analysis investments in like institutions. In all four cases, we used actual data on costs incurred by similar institutions to support these typical programs. These data were provided by a total of eleven institutions with which we have worked closely on gathering assessment data and on using assessment results to improve program planning and decision making.

We constructed all four estimates by means of a specially-designed microcomputer template using the Lotus 1-2-3 Spreadsheet (copyright) program. The template embodies available cost data on eight commercial test and survey

instruments, as well as routines for estimating the costs of designing local test and survey instruments and of administering tests and surveys in classroom and mailed formats, and for estimating overhead costs associated with establishing an assessment office. The template contains on-line instructions for creating cost estimates.

Case 1 - Private Liberal Arts College

Case 1 is a small private liberal arts college with a total enrollment of approximately 1,000 students. The student body is "traditional," more than 95 percent attends full-time, and more than three quarters is in residence, living either in dormitories or in nearby private housing. The curriculum is also traditional, including a recently reinstituted general education core program and a typical list of undergraduate arts and sciences majors. There are no explicitly professional or pre-professional programs, although many students go on to professional or graduate training.

Assessment in this case is concentrated on the gain, or "value-added," of the total college experience, particularly in relation to its general education component. Because of the college's mission, the faculty have decided to administer the ACT-COMP Composite Examination to incoming freshmen and to graduating seniors. They have also opted to make maximum use of the COMP through a consulting visit each year in which ACT staff work with faculty in interpreting scores. The college has found that these visits are an important faculty development tool in addition to the information provided by the examination itself.

The college already participates in the CIRP freshman survey to a limited degree, and will supplement the sample to include the entire estimated freshman class (300 students). At the same time, interest in the involvement of currently enrolled students on campus led to a decision to administer the Pace College Student Experiences Questionnaire (CSEQ) to a selected sample of all students (150) each spring. Finally, the college conducts an alumni study every three years, covering the last three graduating classes. The college plans to develop its own survey but meanwhile is using the ACT-ESS Alumni Survey, which it supplements with 10 locally designed questions.

ACT-COMP testing occurs in classroom settings with dorm counselors serving as proctors. Each student receives an announcement of the test date and is provided with his or her own results after scoring. CIRP and CSEQ surveys are administered in class or through campus mail. The major survey effort is the alumni survey, but the small numbers of actual graduates each year do not entail a major cost. The response rate averages 75 percent for these surveys.

To coordinate the testing program, the college has appointed a junior faculty member in psychology as an assessment director at .35 FTE. She is assigned a one-third time secretary to handle announcements, record survey results, etc. Both personnel costs are shown in the estimate as the full replacement costs for these positions. As noted above, however, these will vary with the need for replacing these individuals in their current functions and with the current market costs of such replacement. Overhead costs are already absorbed by the office of the dean of the faculty to whom the assessment director reports.

Total estimated costs for Case 1 are documented in Table 1.

**TABLE 1: Case 1 - Private Liberal Arts College**

| | |
|---|---|
| Instrument Costs | |
| 300 Freshman General Education Exams (ACT-COMP) | $4,500.00 |
| 150 Senior General Education Exams (ACT-COMP) | 2,250.00 |
| 150 Senior Activity Inventories (ACT-COMP) | 525.00 |
| 300 Freshman Surveys (CIRP) | 415.00 |
| 150 Current Student Surveys (Pace CSEQ) | 337.50 |
| 150 Alumni Surveys (ACT-ESS) | 147.50 |
| Administration Costs | |
| In-Class Test Administration | |
| Proctors, etc. | 342.00 |
| Announcements, etc. | 177.50 |
| Mailed Survey Costs (2 mailings) | 193.62 |
| Overhead/Analysis Costs | |
| ACT Comp Consulting Visit (Fee + Travel) | 1,375.00 |
| CIRP Data Analysis | 150.00 |
| Testing/Measurement Specialist (.35 FTE) | 9,625.00 |
| Secretary/Clerk (.35 FTE) | 5,775.00 |
| Staff Benefits | 3,388.00 |
| TOTAL | $29,201.12 |

## Case 2 - Major Public Research University

The second case is a major public research university with a total enrollment of over 25,000 students, including about 18,000 undergraduates. Faculty make considerable research contributions to their own disciplines and concentrate much of the teaching energy on graduate instruction. Most introductory courses are lecture classes and are partially staffed by graduate teaching assistants. Most undergraduate students attend full-time, and about two-thirds are residential. Attrition rates are significant, but about 65 percent of entering students complete their degrees. Professional schools account for approximately 60 percent of undergraduate enrollment.

Because of its emphasis on professional and pre-professional study, much assessment effort has gone into testing in the major field. Graduates of about 10 programs per year are tested using available standardized test instruments. This year, 450 students are to be tested using a variety of GRE Field Examinations, and 360 students are to be tested using pre-professional examinations such as the National Teacher Examination (NTE) and the AICPA exam. In addition, the institution is evaluating general education using the ACT-COMP Objective Test in a test-retest format for freshmen and seniors. Like Case 1, the institution has budgeted for a consulting/faculty development visit in conjunction with the COMP.

To examine student life, the university intends to design its own survey, using faculty expertise. The survey will be administered to a stratified random sample of currently enrolled students in the spring. Because of the size of the campus and the characteristics of the sample, a mailed format will be used to administer the survey. Similar surveys in the past have obtained a 65 percent response rate.

To coordinate the testing program, the university has staffed an existing student research office with two new staff members, a testing specialist and a secretary. Existing senior staff in the testing office are also used in interpreting test results and in working with individual program faculties on improving curricula. Because many of the fields offered by the university are not now covered by an available standardized senior-level examination, testing center personnel are expected to work with program faculty to design local achievement tests as part of their budgeted assignment. Three such fields are scheduled for test construction this year in geology, archeology and food technology. Approximately $2,500 per test is budgeted for this activity to be paid to participating departments. Other test development costs are expected to be covered through department-level reallocation of faculty staff time already committed to curriculum review.

Table 2 presents total estimated costs for this case.

**TABLE 2: Case 2 - Major Public Research University**

| | |
|---|---|
| Instrument Costs | |
| 2500 Freshman General Education Exams (ACT-COMP Objective Test) | $15,000.00 |
| 1700 Senior General Education Exams (ACT-COMP Objective Test) | 10,200.00 |
| 450 Senior Field Exams (GRE) | 13,050.00 |
| 360 Senior Field Exams | 7,500.00 |
| Development Cost for 3 Local Field Examinations (Professional and Pre-Professional) | 9,270.00 |
| Development Cost for Student Survey | 5,200.00 |
| 2025 Surveys (Producing and Scoring Card) | 518.75 |
| Administration Costs | |
| In-Class Test Administration | |
| Proctors, etc. | 1,826.00 |
| Announcements, etc. | 2,077.00 |
| Mailed Survey Costs (2 mailings) | 1,957.00 |
| Overhead/Analysis Costs | |
| ACT COMP Consulting Visit | 1,375.00 |
| ACT-COMP Data Tape | 20.00 |
| Testing/Measurement Specialist (1 FTE) | 27,500.00 |
| Secretary 16,500.00 | |
| Staff Benefits 9,680.00 | |
| Office Expenses | 8,400.00 |
| TOTAL | $130,073.75 |

Case 3 - Regional Comprehensive University

Case 3 is a public regional comprehensive university enrolling approximately 5,500 students, including 4,500 undergraduates. Like many of its type, the university is a former teachers' college which became a comprehensive university in the early 1970s. In addition to liberal arts disciplines, the university now offers a range of professional subjects through the master's level. These are dominated by education and business, which together enroll about half of the student body. Forty percent of the undergraduate students attend part-time and about two-thirds commute. The university does not currently commit significant resources to academic administration and support, and is proud of its tradition of "low overhead."

As in Case 2, the university seeks to insure that graduating seniors have received adequate training in the major field. Therefore, it has chosen to administer standardized senior examinations annually to the graduates of a fifth of its departments on a rotating basis. This year approximately 120 graduating seniors in fourteen fields will be tested using a variety of instruments. GRE or pre-professional examinations are used where possible. All but three fields in which degrees are offered by the university currently are covered by an existing standardized examination.

The university is also committed to building basic skills, but its emphasis on general education is insufficient to justify the expense of an instrument such as ACT-COMP. Therefore, the faculty has decided to examine "value-added" by using the ACT assessment administered to entering freshmen and at the end of the sophomore year. Of the approximately 1,200 new freshmen each year, about 300 must be given the ACT assessment at university expense. All 800 sophomores are subsequently tested at university expense.

Finally, the university has elected to use a relatively low-cost, standardized survey system -- the ACT Entering Student Survey -- to investigate student opinion and post-graduate success. All entering students are surveyed using the ACT-ESS, and all graduates are surveyed a year after graduation. Finally, a sample of withdrawing students is followed up every other year with the ACT Withdrawing Student Survey. Entering student questionnaires are administered at freshman orientation. Other surveys are administered by mail. All scoring is done by ACT, although the university purchases extra reports and tapes for local analysis. Few local analyses of these data, however, have actually been conducted.

To coordinate testing, the university grants one-third release time to a faculty member in sociology. As before, the full cost of replacement is provided here, although replacement costs using part-time instructors have in practice been less. Work-study students are used to support the survey effort, and graduate students in education are used as test proctors for sophomore and senior examinations.

Costs for Table 3 are itemized in Table 3.

TABLE 3: Case 3 - Regional Comprehensive University

| | |
|---|---|
| **Instrument Costs** | |
| 300 Freshman Tests (ACT Assessment; 900 assumed to have scores on entrance) | $2,100.00 |
| 800 Sophomore Tests (ACT Assessment) | 5,600.00 |
| 1200 Freshman Interest Inventories (ACT Assessment) | 3,000.00 |
| 80 Senior Field Exams (GRE) | 2,320.00 |
| 40 Senior Field Exams (Professional and Pre-Professional) | 940.00 |
| 1200 Entering Student Surveys (ACT-ESS) | 240.00 |
| 350 Non-Returning Student Surveys (ACT-ESS) | 70.00 |
| 650 Alumni Surveys (ACT-ESS) | 130.00 |
| Scoring for 2200 ACT-ESS Instruments | 1,040.00 |
| **Administration Costs** | |
| In-Class Test/Survey Administration | |
| Proctors, etc. | 375.00 |
| Announcements, etc. | 580.00 |
| Mailed Survey Administration (2 mailings) | 1,378.00 |
| **Overhead/Analysis Costs** | |
| ACT-ESS Tape/Reports | 270.00 |
| Testing/Measurement Specialist (.35 FTE) | 9,625.00 |
| Staff Benefits | 2,118.00 |
| Work Study Students | 1,750.00 |
| Office Expenses | 1,250.00 |
| TOTAL | $32,786.00 |

## Case 4 - Mid-Sized Community College

Case 4 is a community college located in a suburb of a major city, enrolling approximately 15,000 students each term. Enrollment consists of about 3,500 baccalaureate transfer students and 5,000 students in various occupational and certificate programs, with the balance enrolled for one or more single courses. About half of the students in baccalaureate transfer and occupational programs attend full-time, many of them at night. All other students are part-time attenders. All students commute to the campus from within a 30-mile radius.

The primary emphasis of assessment at the college has been properly placed upon student follow-up and the assessment of educational goals. All entering program students are surveyed at registration using the NCHEMS/College Board Student Outcomes Information Service (SOIS) Entering Student Questionnaire. In addition, each year those completing a program and those withdrawing from programs are surveyed by mail using SOIS instruments. These surveys achieve approximately 70 percent and 45 percent response rates, respectively. Local questions are added to all SOIS questionnaires, and the Institutional Research Office conducts analyses which link common questions on the three instruments

in order to obtain a composite picture of student reactions to the college experience. All SOIS scoring is done by the College Board, although the college plans to develop its own computer programs to facilitate a more detailed analysis of these instruments.

This year, reacting to statewide concerns about the quality of basic skills education, the college plans to administer the College Level Examination Program general exams in writing and quantitative skills. This will be an expensive effort and is being undertaken somewhat reluctantly. Despite the fact that it was not designed for curriculum evaluation, the CLEP was chosen by faculty as being the most appropriate available instrument to measure general lower-division competence in these areas. It will be given to a sample of 750 second-year program students.

The student follow-up effort was begun several years ago in response to federal Vocational Education Data System (VEDS) requirements. This system mandated student follow-up surveys for graduates of all federally-funded occupational programs. To meet the demand, the college created a half-time data analyst position in the Office of Institutional Research. As assessment has expanded, the responsibility for conducting all studies has remained with Institutional Research; an additional half-time position will be added for survey coordination and to help with administering the CLEP.

Total estimated costs for assessment at Case 4 are presented in Table 4.

**TABLE 4: Case 4 - Mid-Sized Community College**

| | |
|---|---|
| **Instrument Costs** | |
| 750 Sophomore General Skills Exams (CLEP General) | $19,500.00 |
| 4500 Entering Student Surveys (SOIS) | 675.00 |
| 1500 Former Student Surveys (SOIS) | 225.00 |
| 1250 Graduate Follow-Up Surveys (SOIS) | 187.50 |
| Scoring for 7250 SOIS Instruments | 3,490.00 |
| **Administration Costs** | |
| In-Class Test Administration | |
| Proctors, etc. | 325.00 |
| Announcements, etc. | 225.00 |
| Mailed Survey Administration (2 mailings) | 3,693.00 |
| **Overhead/Analysis Costs** | |
| Tapes/Reports of SOIS Surveys | 150.00 |
| Student Survey Coordinator (.5 FTE) | 10,750.00 |
| Staff Benefits | 2,365.00 |
| Office Benefits | 1,500.00 |
| TOTAL | $43,085.50 |

To check the validity of each of these cost estimates, we obtained actual cost data from a total of eleven institutions covering all four of our "case" categories. Total costs for assessment at each of these institutions, of course, vary somewhat from our constructed estimates and from one another. This variance occurs because each institution measures a somewhat different set of outcomes dimensions, and also because the manner in which actual costs are counted and reported are different between institutions. For reasons of confidentiality, we do not report these actual costs. However, when adjusting for total enrollment, none differs by more than 15 percent from our constructed estimates.

## SOME CONCLUDING POINTS

Each of the cases presented above represents a distinctive match between institutional mission and characteristics on the one hand, and a particular choice of assessment instruments and methods on the other. Any cost estimate must be similarly tailored to fit a particular situation. In conclusion, institutions considering implementing a comprehensive assessment program and examining the cost consequences should consider the following points:

- Making full use of existing information about student learning and development can considerably reduce anticipated costs of assessment. As emerging institutional experience makes clear, colleges and universities generally collect considerable information about students; but this information is rarely centrally available. Indeed, no single person or office at the institution may know the full range of what is available. Many individual units may collect data for different purposes. For example, individual departments may collect follow-up information on their own graduates, student service offices may conduct surveys of currently enrolled students, and testing offices may administer a variety standardized tests. A first step in constructing an assessment program is often simply to inventory such data (Ewell 1982).

- Developing an explicit assessment program may reduce cost by focusing analytical and data collection resources and avoiding duplication. Emerging institutional experience has also shown that gathering data on student outcomes can often be inefficient due to its dispersal throughout the institution. Different units develop their own assessment instruments independently, and incur costs in doing so. Furthermore, many studies are one-shot -- designed to answer a particular question or to address a particular, temporary crisis. When the question is answered or the crisis passed, data gathering ceases, only to be begun from scratch when the next question arises. Central coordination of assessment can avoid such hidden costs, and may consequently involve fewer new resources than initially anticipated.

- Assessment programs using multiple data collecting methods may similarly reduce costs by providing mutually reinforcing information. Cognitive testing, for example, is expensive compared to other forms of outcomes data gathering. While there is no substitute for testing

to answer questions concerning student learning in general education or in the major field, much can be learned by supplementing cognitive tests with less expensive kinds of data collection -- for example, student surveys containing self-assessments of growth. If survey information can confirm the results of cognitive tests in the aggregate, expensive testing for purposes of program evaluation may then be undertaken using only small but carefully selected samples of students.

- o Careful tailoring of data collection to fit instructional mission can limit cost. A major potential problem with assessment, as mentioned above, is the implicit assumption that it will "measure everything that moves." Paying close attention to priority instructional and curricular issues in designing an assessment program involves making appropriate choices about what to measure and how to measure it. Each of the cases we constructed, for example, places the primary weight of assessment upon a particular dimension that matches the institution's unique curriculum and mission. Each could have been quite different, and considerably more expensive, if limiting choices had not been exercised.

A final point is that the costs of assessment are in themselves of little importance without knowing the benefits. Many of the 22 institutions involved in the NCHEMS/Kellogg Student Outcomes Project in 1982-85, for example, found that the long-term benefits of assessment information included increases in student recruitment and retention (Ewell 1984). In the long term, such benefits can involve fiscal, as well as strictly educational, rewards. As a result, any assessment program is properly seen as not simply a cost to be incurred, but as an investment in the institution's future: an investment which should be judged in the light of the return that it may bring.

REFERENCES

American College Testing Program. Evaluation Survey Service. Iowa City, Iowa: ACT, 1985.

Astin, A. W.; Green, K. C.; Korn, W. S.; and Maier, M. J. "The American Freshman: National Norms for Fall 1984." Los Angeles, Calif.: Higher Education Research Institute, 1984.

Chancellor's Office, California Community Colleges. "Item Bank: Student and College Outcomes." Improving Community College Evaluation and Planning Project. Sacramento, Calif, Fall 1984.

Dill, Don A. "Mail and Telephone Surveys: The Total Design Method." New York, N.Y.: John Wiley and Sons, 1982.

Ewell, Peter T. "The Self-Regarding Institution: Information for Excellence." Boulder, Colo.: National Center for Higher Education Management Systems, 1984.

Ewell, Peter T. "Information on Student Outcomes: How to Get it and How to Use it." Boulder, Colo.: NCHEMS, 1982.

Forrest, Aubrey. "Increasing Student Competence and Persistence: The Best Case for General Education?" Iowa City, Iowa: ACT, 1982.

Higher Education Research Institute. "Composite Comparative Data College Student Experiences Questionnaire 1979-1982." Los Angeles, Calif.: HERI, 1982.

Lenning, Oscar T. "The Outcomes Structure: An Overview and Procedures for Applying It in Postsecondary Education Institutions." Boulder, Colo.: NCHEMS, 1977.

Mentkowski, Marcia, and Doherty, Austin. "Careering After College: Establishing the Validity of Abilities Learned in College for Later Careering and Professional Performance." Final report to the National Institute for Education. Milwaukee, Wis: Alverno Productions, 1983 (ed. 239,556-239,556).

McKenna, Barbara. "Surveying Your Alumni: Guidelines and 22 Sample Questionnaires." Washington, D.C.: Council for Advancement and Support of Education, 1981.

National Institute of Education, Study Group on the Conditions of Excellence in American Higher Education. "Involvement in Learning: Realizing the Potential of American Higher Education." Washington, D.C.: U.S. Government Printing Office 1984.

Pace, C. Robert, and Associates. "Higher Education Measurement and Evaluation Kit." Los Angeles, Calif.: Laboratory for Research on Higher Education, UCLA, 1975.

APPENDIX

Many commercial instruments are available for assessing student cognitive growth, reactions to college, and experiences after graduation. This appendix lists some of the most commonly used such instruments, together with the direct costs associated with purchasing and administering them.

Institutional assessment programs vary greatly in the outcomes dimensions included. The choice of dimensions (and therefore of particular instruments and emphases) will depend upon the mission, type, and size of the institution, and most of all on its distinctive instructional philosophy. Typical outcomes dimensions and the means commonly chosen to assess them are as follows:

| | |
|---|---|
| 1. General Knowledge/Skills Outcomes: | ACT-COMP |
| eg. Reading/Writing/Math | ACT Assessment |
| Interdisciplinary Knowledge | GRE Comprehensive |
| Ability to Apply Knowledge | CLEP Writing/Math |
| Critical Thinking, etc. | Local Tests |
| 2. Discipline/Field Outcomes: | GRE Field Exams |
| eg. Specific Content Area | CLEP Subject Tests |
| Research Skills | Pre-Professional Tests |
| Professional Competence | Local Tests |
| 3. Student Attitudes/Values: | ACE/UCLA - CIRP |
| eg. Perceptions of Environment | ACT-ESS |
| Self-Assessment of Growth | NCHEMS - SOIS |
| Use of Campus Resources | UCLA/Pace - CSEQ |
| Educational/Career Goals | TEX-SIS |
| | Local Surveys |
| 4. Behavior After College | ACT-ESS |
| eg. Employment History/Success | NCHEMS - SOIS |
| Further Education | TEX-SIS |
| Evaluation of Program | Local Surveys |
| Reasons for Non-Completion | |

Each of these dimensions should be carefully examined in developing a comprehensive assessment program, and costs associated with each component estimated only after an appropriate choice has been made.

Particular instruments and their associated costs are described in the pages that follow.

ACT -- COLLEGE OUTCOME MEASURES PROJECT (COMP)

Description

COMP is a set of examinations designed to assess student learning in General Education. The Composite Exam involves oral and written response by students; other instruments involve written as well as aural and visual materials. Scoring of the composite is by trained faculty. The minimum order is $600.

INSTRUMENT COSTS

| Instrument | Fixed Fee | Unit Cost |
|---|---|---|
| Composite Exam | N/A | $18.00* |
| Objective Test | N/A | 12.00** |
| Activity Inventory | N/A | 4.00*** |

* $15 for 100-499 students
$12 for 500-999 students
$10 for over 1000 students

** $10 for 100-499 students
$ 8 for 500-999 students
$ 6 for over 1000 students

*** $3.50 for 100-499 students
$3.00 for 500-499 students
$2.50 for over 1000 students

ANALYSIS/PROCESSING COSTS

| Service | Fixed Fee | Unit Cost |
|---|---|---|
| Tape of Results | $ 20.00 | $ N/A |
| Consulting Visit | $600.00 | $ N/A |
| + Travel Expenses | | |

## ACT ASSESSMENT (COLLEGE ENTRANCE EXAM)

### Description

Designed as a measure of general knowledge and skills in preparation for a college curriculum. Generally used as an entrance examination, but has been used in test-retest format to assess "value-added," or as a measure of general knowledge. Interest Inventory is a student survey designed to be administered with the Assessment.

INSTRUMENT COSTS

| Instrument | Fixed Fee | Unit Cost |
|---|---|---|
| ACT Assessment | N/A | $ 7.00 |
| Interest Inventory | N/A | $ 2.50 |

* Note: Instrument costs include processing/analysis.

ANALYSIS/PROCESSING COSTS

Costs cited above include processing and reporting.

ETS -- GRADUATE RECORD EXAMINATION PROGRAM

Description

Comprehensive Exams (Verbal/Quantitative) designed to assess general readiness for graduate work. Discipline/Field Exams in 30 subject areas designed to assess mastery of field content for graduate study.

INSTRUMENT COSTS

| Instrument | Fixed Fee | Unit Cost |
|---|---|---|
| Comprehensive Exam | N/A | $ 29.00 |
| Specific Field Exam | N/A | $ 29.00 |

* Note: Instrument costs include processing/analysis.

ANALYSIS/PROCESSING COSTS

Costs cited above include processing and reporting.

ETS -- COLLEGE LEVEL EXAMINATION PROGRAM (CLEP)

Description

Designed to assess knowledge of particular course/subject areas of college-level work. General Exams assess overall college-level skills such as mathematics and writing. Specific exams in 30 fields test mastery of content in particular course subject areas.

INSTRUMENT COSTS

| Instrument | Fixed Fee | Unit Cost |
|---|---|---|
| General Exams | N/A | $ c. 30.00 |
| Course/Subject Exams | N/A | $ c. 30.00 |

* Note: Instrument costs include processing/analysis.

ANALYSIS/PROCESSING COSTS

Costs cited above include processing and reporting.

ACE-UCLA -- COOPERATIVE INSTITUTIONAL RESEARCH PROGRAM (CIRP)

Description

A survey of student attitudes, goals, and background including self-assessments of academic abilities and college choices. The freshman survey is national and can be supplemented at any college. A limited follow-up survey is also conducted nationally. Instrument costs include processing/analysis

INSTRUMENT COSTS

| Instrument | Fixed Fee | Unit Cost |
|---|---|---|
| Freshman Survey | 175.00 | $ 0.80 |
| Follow-up Survey | 300.00 | $ 2.25 |

ANALYSIS/PROCESSING COSTS

| Service | Fixed Fee | Unit Cost |
|---|---|---|
| Analysis Crosstabs | 150.00 | $ N/A |
| Tape of Results | 60.00 | $ 0.06 |

ACT -- EVALUATION SURVEY SERVICE (ESS)

Description

Student attitudinal and perceptual surveys designed to assess reactions to college, self-assessments of growth, and evaluations of services provided. Follow-up surveys include employment, graduate school attendance, and reactions to the college program.

INSTRUMENT COSTS

| Instrument | Fixed Fee | Unit Cost |
|---|---|---|
| Entering Student | N/A | $ 0.20 |
| Non-Returning Student | N/A | $ 0.20 |
| Alumni Survey | N/A | $ 0.20 |

ANALYSIS/PROCESSING COSTS

| Service | Fixed Fee | Unit Cost |
|---|---|---|
| Instrument Scoring | 50.00 | $ 0.45 |
| Tape of Results | 40.00 | $ N/A |
| Comparative Report | 50.00 | $ N/A |

## NCHEMS/CEEB -- STUDENT OUTCOMES INFORMATION SERVICE (SOIS)

### Description

Student attitudinal survey instruments. Items include student background, college choice, self-assessments of gain, perceptions of services, and post-enrollment information such as employment/further education.

INSTRUMENT COSTS

| Instrument | Fixed Fee | Unit Cost |
|---|---|---|
| Entering Student | N/A | $ 0.15 |
| Continuing Student | N/A | $ 0.15 |
| Former Student | N/A | $ 0.15 |
| Graduate (2 instruments) | N/A | $ 0.15 |

ANALYSIS/PROCESSING COSTS

| Service | Fixed Fee | Unit Cost |
|---|---|---|
| Keypunch/Processing | 100.00 | $ 0.44 |
| Tape of Results | 50.00 | $ N/A |

## UCLA/Pace -- COLLEGE STUDENT EXPERIENCES QUESTIONNAIRE (CSEQ)

### Description

Student survey instruments designed to assess student quality of effort. Includes activity inventory, involvement with the campus, and perceptions of the college environment.

INSTRUMENT COSTS

| Instrument | Fixed Fee | Unit Cost |
|---|---|---|
| CSEQ | N/A | $ 0.25 |

ANALYSIS/PROCESSING COSTS

| Service | Fixed Fee | Unit Cost |
|---|---|---|
| Analysis/Processing/Report | 150.00 | $ 1.00 |
| Consulting Visit +Travel Expenses | 600.00 | $ |

PART THREE: Additional Resources

## HIGHLIGHTS OF STATE POLICIES

This represents a partial listing of state action in assessment programs to measure what students learn in college. For more detailed information, the reader is referred to "College Quality: Measuring What Students Learn in Higher Education."

Study Commissions

| | |
|---|---|
| Arkansas - | Released the Quality in Higher Education report (1984) recommending, but not requiring, basic skills and rising junior exams; institutions have been asked to implement campus assessment plans. |
| Colorado - | The higher education reorganization plan requires state institutions of higher education to development assessment procedures. States not in compliance by 1990 will lose 2 percent of their appropriation. |
| Louisiana - | Board of Regents has required state postsecondary institutions to assess student learning in relation to the newly revised general education curriculum. |
| Maryland - | The development of a general education proficiency test has been recommended. |
| Minnesota - | The Minnesota Higher Education Coordinating Board sponsored a Task Force on Definitions of College - Level Skills, which recommended reading, writing, and mathematices skills that should be expected of students entering degree programs in all Minnesota postsecondary institutions. |
| Utah - | Board of Regents strongly encourage postsecondary institutions to assess student learning. Institutions must submit annual reports on student outcomes. |
| Virginia - | A panel recommended decentralized postsecondary institutional assessment practices. |

## Entry-Level Testing

| | |
|---|---|
| Alabama - | State college entrants undergo placement testing; a "Four Years Plus" program serves students with severe deficiencies in the basic skills. |
| New Jersey - | The New Jersey Basic Skills Placement Test (NJBSPT) is given to all incoming students. Non-credit remediation is required for students deficient in basic skills. |
| Ohio - | Early Mathematics Placement Program; Early English Composition Assessment Program involve college placement tests administered to high school juniors. Those found to need remedial work may make it up during the senior year, rather than in college remedial classes. |

## Basic Skills Proficiency for Graduation

| | |
|---|---|
| California - | State university graduates must demonstrate English proficiency prior to graduation. |
| Florida - | College sophomores must pass the College Level Academic Skills Test (CLAST) in order to obtain an associate degree or enroll in upper division work. |
| Georgia - | Students must pass the Regents testing program to demonstrate writing and reading skills before graduation. |

## Value-added Assessment

| | |
|---|---|
| Kentucky - | The Council of Higher Education recommended value-added assessment on a pilot basis (1985). |
| Missouri - | Northeast Missouri State University tests freshmen students at entry with ACT or ACT COMP; students are re-tested at end of the sophomore year. Senior take subject-matter specialty exams (such as GRE) prior to graduation. Student, alumni, employers are surveyed to further assess the institutional performance. |
| South Dakota - | The regents have adapted the NMSU model to the South Dakota higher education system. |

Performance-Based Funding

| | |
|---|---|
| Massachusetts - | Funds Centers of Excellence. |
| New Jersey - | Governor's challenge funds target state institutions that demonstrate improvement. |
| Ohio - | Has adopted five grant programs to improve higher education. |
| South Dakota - | Has begun a performance funding program. |
| Tennessee - | State colleges and universities receive up to a five percent bonus for meeting five performance criteria, including student performance on general education measures, subject matter mastery, and evidence of student satisfaction with their college. |

## ADDITIONAL SOURCES

The reader who wishes additional research information and resource material is referred to the sources listed below. The clearinghouse is funded and coordinated under the aupices of the U.S. Department of Education's Office of Educational Research and Improvement.

ERIC Clearinghouse on Teacher Education

The Clearinghouse on Elementary and Early Childhood Education has established a data base of relevant articles, papers, publications, and materials on topics relating to college and university conditions, problems, programs, and students, curricular and instructional programs, and institutional research at the college or university level. Other issues covered include professional continuing education, collegiate computer-assisted learning and management, graduate education, university extension programs, teaching-learning, legal issues and legislation, planning, governance, finance, evaluation, interinstitutional arrangements, management of institutions of higher education, business or industry educational programs leading to a degree, federal programs, and professional education.

George Washington University
One Dupont Circle, N.W., Suite 630
Washington, DC 20036
(202) 296-3597

Center for Improving Postsecondary Learning and Teaching

The center focuses its research, development, and dissemination activities on five aspects of college learning environments that affect learner outcomes: classroom learning and teacher strategies, curricular structure and integration, faculty attitudes and teaching behaviors, organizational practices, and use of emerging information technology. While recognizing multiple student outcomes of college such as cognitive development, personal development, and career development, the center emphasizes cognitive development of undergraduate students in colleges that concentrate on teaching as their primary mission. This emphasis was chosen because the recent dramatic progress of research in cognition holds great promise for improving learning and teaching. Major projects of interest to readers include programs on instructional processes and educational outcomes, curricular integration and education outcomes, teaching and learning, and faculty as a key resource.

School of Education
University of Michigan
Ann Arbor, MI 48109
(313) 764-9472

Other relevant information may be solicited from the Center for Postsecondary Governance and Finance.

The center seeks to improve the effectiveness, efficiency, and equity of teaching, research, and public service in postsecondary education through programs of policy research and dissemination of information designed to improve governance, management, and finance practices. The concerns of post-secondary education will be pursued at the institutional, interinstitutional, state, and federal levels. Aims include: 1) promoting learning through teaching, 2) creating new knowledge through research and scholarship, and 3) disseminating knowledge and providing assistance through public service activities.

University of Maryland
Institute for Research in Higher and Adult Education
College of Education
College Park, MD 20742

Higher education issues are one area of interest for the Center on Education and Employment. The center seeks to determine the implications for education and training of important changes occurring in employment and the economy. Education and training provided through formal systems of schoolling and through employers or other kinds of institutions is a focus of the center. The quality and content, timing and place of education and training in relation to the worklives of individuals and the performance of the economy are emphasized.

Teachers College, Columbia University
Box 174
New York, NY 10027
(212) 678-3091